

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/009959

International filing date: 23 March 2005 (23.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/555,765
Filing date: 24 March 2004 (24.03.2004)

Date of receipt at the International Bureau: 25 April 2005 (25.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1310056

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

April 18, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/555,765

FILING DATE: *March 24, 2004*

RELATED PCT APPLICATION NUMBER: *PCT/US05/09959*



Certified by

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

032404

17169 U.S. PTO

3

Please type a plus sign (+) inside this box PTO/SB/16 (5-03)
Approved for use through 4/30/2003. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

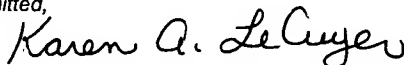
PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

INVENTOR(S)					
Given Name (first and middle [if any])		Family Name or Surname		Residence (City and either State or Foreign Country)	
Mingjun Yongnian Wengang		Huang Sun Yang		Potomac, Maryland Hamden, Connecticut Cheshire, Connecticut	
<input type="checkbox"/> Additional inventors are being named on the _____ separately numbered sheets attached hereto					
TITLE OF THE INVENTION (280 characters max)					
QUANTITATIVE ASSAY FOR DETECTION OF NEWLY SYNTHESIZED RNA IN A CELL-FREE SYSTEM AND IDENTIFICATION OF RNA SYNTHESIS INHIBITORS					
Direct all correspondence to: CORRESPONDENCE ADDRESS					
<input checked="" type="checkbox"/> Customer Number		23413		<div>Place Customer Number Bar Code Label here</div>	
OR Type Customer Number here					
<input type="checkbox"/> Firm or Individual Name					
Address					
Address					
City		State	ZIP		
Country		Telephone	Fax		
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/> Specification		Number of Pages	37	<input type="checkbox"/> CD(s), Number	
<input checked="" type="checkbox"/> Drawing(s)		Number of Sheets	2	<input checked="" type="checkbox"/> Other (specify)	
<input type="checkbox"/> Application Data Sheet. See 37 CFR 1.76		Informal DNA Sequence Listing			
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)					
<input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27.				FILING FEE AMOUNT (\$)	
<input type="checkbox"/> A check or money order is enclosed to cover the filing fees					
<input checked="" type="checkbox"/> The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number		06-1130		\$80.00	
<input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.					
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input checked="" type="checkbox"/> No.					
<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: _____					

Respectfully submitted,

SIGNATURE



Date 3/24/04

TYPED or PRINTED NAME

Karen A. LeCuyer

REGISTRATION NO.

51,928

(if appropriate)

Docket Number:

API-0005

TELEPHONE

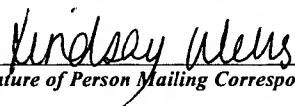
860.286.2929

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is used by the public to file (and by the PTO to process) a provisional application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the complete provisional application to the PTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop Provisional Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

P19SMALL/REV05

CERTIFICATE OF MAILING BY "EXPRESS MAIL" (37 CFR 1.10) Applicant(s): HUANG et al			Docket No. API-0005
Serial No. NYA	Filing Date NYA	Examiner NYA	Group Art Unit NYA
Invention: QUANTITATIVE ASSAY FOR DETECTION OF NEWLY SYNTHESIZED RNA IN A CELL-FREE SYSTEM AND IDENTIFICATION OF RNA SYNTHESIS INHIBITORS			
<p>I hereby certify that this <u>Provisional Patent Application</u> <i>(Identify type of correspondence)</i></p> <p>is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 in an envelope addressed to: Director of the United States Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450 on <u>March 24, 2004</u> <i>(Date)</i></p> <p><u>Lindsay Wells</u> <i>(Typed or Printed Name of Person Mailing Correspondence)</i></p> <p><u></u> <i>(Signature of Person Mailing Correspondence)</i></p> <p><u>EV 402807847 US</u> <i>("Express Mail" Mailing Label Number)</i></p>			
<p>Note: Each paper must have its own certificate of mailing.</p>			

PROVISIONAL PATENT APPLICATION IN THE UNITED STATES PATENT AND
TRADEMARK OFFICE

TITLE: QUANTITATIVE ASSAY FOR DETECTION OF NEWLY SYNTHESIZED RNA IN
A CELL-FREE SYSTEM AND IDENTIFICATION OF RNA SYNTHESIS INHIBITORS

INVENTORS:

HUANG, MINGJUN (11621 GREENLANE DRIVE, POTOMAC, MARYLAND; CITIZENSHIP: USA)

SUN, YONGNIAN, (80 HOWARD DR., HAMDEN, CT 06514; CITIZENSHIP: P.R. CHINA)

YANG, WENGANG (114 OLD FARMS RD, CHESHIRE, CT 06410; CITIZENSHIP: P.R.CHINA)

ASSIGNED TO:

ACHILLION PHARMACEUTICALS

300 GEORGE ST.

NEW HAVEN, CT 06510

QUANTITATIVE ASSAY FOR DETECTION OF NEWLY SYNTHESIZED RNA IN A CELL-FREE SYSTEM AND IDENTIFICATION OF RNA SYNTHESIS INHIBITORS

BACKGROUND

[0001] HCV is one of the most prevalent causes of chronic liver disease in the United States, which accounts for about 15 percent of acute viral hepatitis, 60 to 70 percent of chronic hepatitis, and up to 50 percent of cirrhosis, end-stage liver disease, and liver cancer. Almost 4 million Americans, or 1.8 percent of the U.S. population, have antibodies to HCV (i.e., anti-HCV antibodies), indicating ongoing or previous infection with the virus. Hepatitis C causes an estimated 8,000 to 10,000 deaths annually in the United States. While the acute phase of HCV infection is usually associated with mild symptoms, some evidence suggests that only about 15% to 20% of the infected people will clear HCV.

[0002] HCV is a small, enveloped, single-stranded positive RNA virus in the Flaviviridae family. The genome is approximately 10,000 nucleotides and encodes a single polyprotein of about 3,000 amino acids. All of the protein products of HCV are produced by proteolytic cleavage of the polyprotein, carried out by one of three proteases: the host signal peptidase, the viral self-cleaving metalloproteinase, NS2, and the viral serine protease NS3/4A. The combined action of these enzymes produces the structural proteins (C, E1 and E2) and non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) that are required for replication and packaging of the viral genomic RNA. NS5B is the viral RNA-dependent RNA polymerase (RDRP) that is responsible for the conversion of the input genomic RNA into a minus strand copy (complimentary RNA, or cRNA); the cRNA then serves as a template for transcription by NS5B of the positive sense genomic/messenger RNA. The HCV replicase is the complex of proteins that are necessary for the accurate and efficient synthesis of viral replicon RNA.

[0003] Currently, the only effective therapy against HCV is alpha-interferon, which reduces the amount of virus in the liver and blood (e.g., viral load) in only a small proportion of infected patients. Standard forms of interferon, however, are now being replaced by pegylated interferons (peginterferons), alpha interferons that have been modified chemically by the addition of a large inert molecule of polyethylene glycol. At the present time, the optimal regimen including interferon appears to be a 24- or 48-week course of a combination of pegylated alpha interferon and the nucleoside ribavirin, an oral antiviral agent that has activity

against a broad range of viruses. Nonetheless, response rates to the combination interferon/ribavirin therapy may be moderate for certain HCV genotypes, i.e., a response rate of about 50% to about 60%, although response rates for selected genotypes of HCV (notably genotypes 2 and 3) are typically higher. Another drawback to the current therapy is that there are often significant adverse side effects associated with each of these agents including, for example, flu-like symptoms; bone marrow suppressive effects; neuropsychiatric effects such as marked irritability, anxiety, personality changes, depression, and even suicide or acute psychosis; histamine-like side effects; and anemia.

[0004] Taken together, the preceding facts indicate a significant need for effective small molecule inhibitors of HCV replication that do not suffer from the above-mentioned drawbacks. A particularly useful class of inhibitors of HCV, as well as other positive strand RNA viruses, is inhibitors of viral RNA synthesis.

[0005] While accurate and efficient assays for identifying HCV RNA synthesis inhibitors may be useful tools for identifying effective small molecule HCV therapeutics, no such system has been developed. An *in vitro* replication assay using recombinant NS5B polymerase has been reported. However, in this system, the purified form of NS5B polymerase lacked template specificity and produced various lengths of RNA products. These phenomena are very different from HCV RNA replication *in vivo*. To better reflect the HCV RNA replication process in the cell, a cell-free HCV replication system was established using whole cell lysates or membrane fractions of cells expressing the HCV replicon. In this cell-free system, radioactive P³²-UTP or P³²-CTP was used to label newly synthesized HCV RNA, and then the reaction products were resolved by gel electrophoresis, followed by autoradiography. Because these assays require gel electrophoresis to separate the full length HCV RNA from other RNA molecules, the results are difficult to quantify, often inaccurate, and poorly reproducible. Additionally, while it appears to be clear that RNA elongation occurs in this cell-free system, there is no convincing evidence that *de novo* RNA initiation occurs in this system.

SUMMARY

[0006] An assay for detecting newly synthesized RNA in a cell-free system, and for identifying compounds that inhibit positive strand RNA viruses such as the Hepatitis C Virus (HCV) is provided herein.

[0007] Provided herein is an assay for detecting newly synthesized RNA from a positive strand RNA virus, such as HCV virus, which is more efficient and quantitative than previously reported assays. This assay is useful for identifying inhibitors of RNA synthesis of a positive strand RNA virus, including inhibitors of HCV RNA synthesis. Certain embodiments of the assay include methods for distinguishing RNA synthesis initiation inhibitors from RNA elongation inhibitors.

[0008] Provided herein is a method for determining whether a test compound inhibits RNA synthesis of a positive strand RNA virus. The method comprises:

contacting an isolated replicase complex for the positive strand RNA virus comprising a viral replicon template RNA for the positive strand RNA virus, a labeled nucleotide analog, and the test compound, under conditions sufficient for *in vitro* RNA synthesis, to form a newly synthesized RNA population comprising the labeled nucleotide analog;

detecting the newly synthesized RNA population comprising the labeled nucleotide analog;

quantitating the newly synthesized RNA population comprising the labeled nucleotide analog to provide a test RNA amount; and

comparing the test RNA amount with a control RNA amount of a newly synthesized RNA population comprising the labeled nucleotide analog produced in the absence of the test compound, wherein a decrease in the test RNA amount compared to the control RNA amount indicates that the test compound inhibits RNA synthesis of the positive strand RNA virus.

[0009] Further provided herein is a method for quantitating newly initiated RNA of a positive strand RNA virus comprising:

contacting an isolated replicase complex for the positive strand RNA virus comprising a viral replicon template RNA for the positive strand RNA virus, and a labeled nucleotide analog, under conditions sufficient for *in vitro* RNA synthesis, to form a newly synthesized RNA population comprising the labeled nucleotide analog;

hybridizing a probe and the newly synthesized RNA population comprising the labeled nucleotide analog, under stringent hybridization conditions, wherein the probe is complementary

to at least a portion of a transcription initiation region of the newly synthesized RNA population;
digesting unhybridized, single-stranded RNA with a single-strand specific ribonuclease to form a protected RNA population;

detecting a protected RNA population comprising the labeled nucleotide analog; and
quantitating the protected RNA population comprising the labeled nucleotide analog.

[0010] Also provided is a method for determining whether a test compound is an RNA synthesis initiation inhibitor of a positive strand RNA virus comprising:

contacting an isolated replicase complex for the positive strand RNA virus comprising a viral replicon template RNA for the positive strand RNA virus, a labeled nucleotide analog, and the test compound, under conditions sufficient for *in vitro* RNA synthesis, to form a newly synthesized RNA population comprising the labeled nucleotide analog;

hybridizing a probe and the newly synthesized RNA population comprising the labeled nucleotide analog, under stringent hybridization conditions, wherein the probe is complementary to at least a portion of an initiation region of the newly synthesized RNA population;

digesting unhybridized, single-stranded RNA with a single-strand specific ribonuclease to form a protected RNA population;

detecting a protected RNA population comprising the labeled nucleotide analog;

quantitating the protected RNA population comprising the labeled nucleotide analog to provide a test RNA amount; and

comparing the test RNA amount with a control RNA amount of protected RNA comprising the labeled nucleotide analog but produced in the absence of the test compound, wherein a decrease in the test RNA amount compared to the control RNA amount indicates that the test compound inhibits RNA synthesis initiation of the positive strand RNA virus.

BRIEF DESCRIPTION OF THE FIGURES

[0011] Figure 1 shows a schematic of the inhibition of HCV replicon replication by transcription initiation and elongation inhibitors.

[0012] Figure 2 shows an autoradiogram illustrating use of an RNA synthesis inhibitor to demonstrate *de novo* initiation of RNA synthesis in isolated replication complexes.

DETAILED DESCRIPTION

[0013] Prior to setting forth the invention in detail, it may be helpful to provide definitions of certain terms to be used herein. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs.

TERMINOLOGY AND MOLECULAR DESCRIPTION

[0014] A “replicon” as used herein includes a genetic element, for example, a plasmid, cosmid, bacmid, phage or virus that is capable of replication largely under its own control. A replicon may be either RNA or DNA, and may be single or double-stranded. A preferred replicon is a replicon from a single-stranded positive strand RNA virus such as a virus of the Picornaviridae family, the Calciviride family, the Togaviridae family, the Coronaviridae family, or the Flaviviridae family. Viruses of the Flaviviridae family include, for example, Hepatitis C Virus, West Nile Virus, Dengue Virus, Kunjuin Virus, Yellow Fever Virus, Bovine Viral Diarrhea Virus, Tick Born Encephalitis Virus, Japanese Encephalitis Virus, and Venezuelan Equine Encephalitis Virus.

[0015] In positive strand RNA viruses, replication is performed by a multi-protein-RNA complex called a “replicase complex”. As used herein, a “replicase complex” is an active complex of polypeptides and RNA which is capable of complete and accurate viral replicon RNA synthesis under cell-free conditions suitable for viral RNA replication. By complete and accurate viral replicon RNA synthesis, it is meant that the replicase complex is capable of producing full-length viral replicon RNAs. In addition, the isolated replicase complex should show specificity for replication of the replicase RNA of the corresponding positive strand RNA virus. An isolated replicase complex preferably comprises a viral replicon template RNA for the positive strand RNA virus. An “isolated replicase complex” is a replicase complex which has been removed from its cellular environment, such as a cell expressing a viral replicon RNA.

“Isolated replicase complex” includes the membrane fraction of a cell expressing the viral replicase RNA. The isolated replicase complex may be separated from the cell nucleus, chromosomal DNA, and cytoplasmic materials, for example. In addition, an isolated replicase complex may comprise one or more polypeptides expressed from a recombinant expression system, so long as the complex is capable of complete and accurate viral replicon RNA synthesis. The replicase complex of HCV, for example, includes the NS5B protein which has RNA-dependent RNA polymerase activity, and other protein factors

[0016] “Nucleic acid” or a “nucleic acid molecule” refers to a DNA or RNA molecule, either single or double-stranded and, if single-stranded, the molecule of its complementary sequence in either linear or circular form. A sequence or structure of a particular nucleic acid molecule can be described according to the normal convention of providing the sequence in the 5' to 3' direction.

[0017] The term “isolated nucleic acid molecule” includes nucleic acid molecules that are separated from an intact cellular environment. An “isolated nucleic acid molecule” may be, for example, a template RNA that is separated from the cell nucleus, chromosomal DNA, and other cellular materials which are not membrane-associated. Moreover, an “isolated” nucleic acid molecule, such as a viral replicon template RNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. In some embodiments, an isolated viral replicon template RNA may be purified as a portion of a membrane fraction of a cell expressing a viral replicon RNA. Such a membrane fraction may also comprise isolated replicase complexes. Substantially free of other cellular material includes, for example, a cellular fraction such as, for example, a membrane bound fraction. By substantially free of other cellular material, it is meant that an isolated nucleic acid molecule may be greater than or equal to about 70%, about 75%, about 80%, about 85%, about 90%, about 95% or about 99% free of unwanted cellular materials, such as, for example, components of cellular fractions other than the membrane-bound fraction.

[0018] In some embodiments, an “isolated” nucleic acid may be free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and/or 3' ends of the nucleic acid) in the genomic DNA or RNA of the organism from which the nucleic acid is derived. For example, the isolated nucleic acid molecule can contain less than about 5 kb, about 4 kb, about 3 kb, about 2 kb, about 1 kb, about 0.5 kb or about 0.1 kb of 5' and/or 3' nucleotide sequences

which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. In this sense, an isolated nucleic acid may be, for example, a DNA vector encoding a viral replicon RNA which has been purified by standard DNA purification methods.

[0019] “Natural allelic variants”, “mutants” and “derivatives” of particular sequences of nucleic acids refer to nucleic acid sequences that are closely related to a particular sequence but which may possess, either naturally or by design, changes in sequence or structure. By closely related, it is meant that greater than or equal to about 75%, but often, greater than or equal to about 90%, of the nucleotides of the sequence match over the defined length of the nucleic acid sequence. Changes or differences in nucleotide sequence between closely related nucleic acid sequences may represent nucleotide changes in the sequence that arise during the course of normal replication or duplication in nature of the particular nucleic acid sequence. Other changes may be specifically designed and introduced into the sequence for specific purposes, such as to change an amino acid codon or sequence in a regulatory region of the nucleic acid. Such specific changes may be made *in vitro* using a variety of mutagenesis techniques or produced in a host organism placed under particular selection conditions that induce or select for the changes. Such sequence variants generated specifically may be referred to as “mutants” or “derivatives” of the original sequence.

[0020] Different “variants” including “natural allelic variants” of, for example, the HCV genome exist in nature. These variants may be alleles characterized by differences in the nucleotide sequences of the gene coding for a protein, or may involve different RNA processing or post-translational modifications. The skilled person can produce variants having single or multiple amino acid substitutions, deletions, additions or replacements. These variants may include inter alia: a) variants in which one or more amino acids residues are substituted with conservative or non-conservative amino acids, b) variants in which one or more amino acids are added, and c) variants in which one or more amino acids include a substituent group.

[0021] “Operatively linked” refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. An expression control sequence operatively linked to a coding sequence is ligated such that expression of the coding sequence is achieved under conditions compatible with the expression control sequences. As used herein, the term “expression control sequences” refers to nucleic acid sequences that regulate the expression of a nucleic acid sequence to which it is operatively

linked. Expression control sequences are operatively linked to a nucleic acid sequence when the expression control sequences control and regulate the transcription and, as appropriate, translation of the nucleic acid sequence. Thus, expression control sequences can include appropriate promoters, enhancers, transcription terminators, a start codon (i.e., atg) in front of a protein-encoding gene, splicing signals for introns (if introns are present), maintenance of the correct reading frame of that gene to permit proper translation of the mRNA, and stop codons. The term “control sequences” is intended to include, at a minimum, components whose presence can influence expression, and can also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences. Expression control sequences can include a promoter. By “promoter” is meant minimal sequence sufficient to direct transcription. Also included are those promoter elements which are sufficient to render promoter-dependent gene expression controllable for cell-type specific, tissue-specific, or inducible by external signals or agents; such elements may be located in the 5' or 3' regions of the gene. Both constitutive and inducible promoters are included.

[0022] The term “probe” as used herein refers to an oligonucleotide or polynucleotide comprising either RNA or DNA, whether occurring naturally as in a purified restriction enzyme digest or produced synthetically, which is capable of annealing with or specifically hybridizing to a nucleic acid with sequences complementary to the probe. A probe may be, for example, a single-stranded RNA transcribed *in vitro* from a DNA template. A probe may be either single-stranded or double-stranded, but is preferably single-stranded. The exact length of the probe will depend upon many factors, including the temperature, the source of the probe, and use. The probes are selected to be “substantially” complementary to a strand of a particular target nucleic acid sequence. This means that the probes are sufficiently complementary so as to be able to “specifically hybridize” or anneal with their respective target strands under a set of pre-determined conditions. The term specifically hybridize means that the probe has a greater probability of hybridizing to its target sequence than other non-target sequences.

[0023] Therefore, the probe sequence need not reflect the exact complementary sequence of the target. For example, a non-complementary nucleotide fragment may be attached to the 5' or 3' end of the probe, with the remainder of the probe sequence being complementary to the target strand. A small (e.g., less than or equal to about 10%) of mismatched nucleotides in the primer and the target sequence is also permissible.

[0024] The term “oligonucleotide” is defined as a nucleic acid molecule comprising two or more ribonucleotides or deoxyribonucleotides, preferably more than three. The exact size of the oligonucleotide will depend on various factors and on the particular application and use of the oligonucleotide. Preferably, an oligonucleotide comprises less than 100 nucleotides, more preferably less than or equal to about 50 nucleotides, and most preferably less than or equal to about 30 nucleotides.

[0025] A “polynucleotide” probe or primer preferably comprises less than or equal to about 1000 nucleotides, more preferably less than or equal to about 800 nucleotides, and most preferably less than or equal to about 500 nucleotides. A polynucleotide primer also preferably comprises greater than or equal to about 100 nucleotides, more preferably greater than or equal to about 150 nucleotides, and most preferably greater than or equal to about 200 nucleotides.

[0026] “Stringency conditions” for hybridization is a term of art which refers to the incubation and wash conditions, e.g., conditions of temperature and buffer concentration, which permit hybridization of a first nucleic acid to a second nucleic acid; the first nucleic acid may be perfectly (i.e., 100%) complementary to the second, or the first and second may share some degree of complementarity which is less than perfect (e.g., 70%, 75%, 85%, 95%, 98%). For example, certain high stringency conditions can be used which distinguish perfectly complementary nucleic acids from those of less complementarity.

[0027] “High stringency conditions”, “moderate stringency conditions” and “low stringency conditions” for nucleic acid hybridizations are explained in *Current Protocols in Molecular Biology* (Ausubel, F.M. *et al.*, “*Current Protocols in Molecular Biology*”, John Wiley & Sons, (1998)). The exact conditions which determine the stringency of hybridization depend not only on ionic strength (e.g., 0.2X SSC, 0.1X SSC), temperature (e.g., room temperature, 42°C, 68°C) and the concentration of destabilizing agents such as formamide or denaturing agents such as SDS, but also on factors such as the length of the nucleic acid sequence, base composition, percent mismatch between hybridizing sequences and the frequency of occurrence of subsets of that sequence within other non-identical sequences. SSC buffer is 150 mM sodium chloride and 15 mM sodium citrate, pH 7.0. High, moderate or low stringency conditions can be determined empirically.

[0028] By varying stringency conditions for hybridization from a level of stringency at which no hybridization occurs to a level at which hybridization is first observed, conditions which will allow a given sequence to hybridize (e.g., selectively) with the most similar

sequences in the sample can be determined.

[0029] Exemplary conditions are described in Krause, M.H. and S.A. Aaronson, *Methods in Enzymology*, 200:546-556 (1991). Also, in, Ausubel, *et al.*, “*Current Protocols in Molecular Biology*”, John Wiley & Sons, (1998), which describes the determination of washing conditions for moderate or low stringency conditions. Washing is the step in which conditions are usually set so as to determine a minimum level of complementarity of the hybrids. Starting from the lowest temperature at which only homologous hybridization occurs, each °C by which the final wash temperature is reduced (holding SSC concentration constant) allows an increase by about 1% in the maximum extent of mismatching among the sequences that hybridize. Generally, doubling the concentration of SSC results in an increase in T_m of about 17°C. Using these guidelines, the washing temperature can be determined empirically for high, moderate or low stringency, depending on the level of mismatch sought.

[0030] For example, a low stringency wash can comprise washing in a solution containing 0.2X SSC/0.1% sodium dodecyl sulfate (SDS) for 10 minutes at room temperature; a moderate stringency wash can comprise washing in a prewarmed solution (42°C) solution containing 0.2X SSC/0.1% SDS for 15 minutes at 42°C; and a high stringency wash can comprise washing in prewarmed (68°C) solution containing 0.1X SSC/0.1% SDS for 15 minutes at 68°C. Furthermore, washes can be performed repeatedly or sequentially to obtain a desired result as known in the art. Equivalent conditions can be determined by varying one or more of the parameters given as an example, as known in the art, while maintaining a similar degree of identity or similarity between the target nucleic acid molecule and the primer or probe used.

[0031] A “test compound” as defined herein refers to a chemical, nucleic acid, polypeptide, amino acid, or other compound which is to be tested. Examples of test compounds include, but are not limited to, drug candidates, such as derived from arrays of small molecules generated through general combinatorial chemistry, as well as any other substances thought to have potential biological activity.

DESCRIPTION OF THE ASSAYS

[0032] Provided herein is a method for determining whether a test compound inhibits RNA synthesis of a positive strand RNA virus. The method comprises:

contacting an isolated replicase complex for the positive strand RNA virus comprising a

viral replicon template RNA for the positive strand RNA virus, a labeled nucleotide analog, and the test compound, under conditions sufficient for *in vitro* RNA synthesis, to form a newly synthesized RNA population comprising the labeled nucleotide analog;

detecting the newly synthesized RNA population comprising the labeled nucleotide analog;

quantitating the newly synthesized RNA population comprising the labeled nucleotide analog to provide a test RNA amount; and

comparing the test RNA amount with a control RNA amount of a newly synthesized RNA population comprising the labeled nucleotide analog produced in the absence of the test compound, wherein a decrease in the test RNA amount compared to the control RNA amount indicates that the test compound inhibits RNA synthesis of the positive strand RNA virus.

[0033] Further provided herein is a method for quantitating newly initiated RNA of a positive strand RNA virus comprising:

contacting an isolated replicase complex for the positive strand RNA virus comprising a viral replicon template RNA for the positive strand RNA virus, and a labeled nucleotide analog, under conditions sufficient for *in vitro* RNA synthesis, to form a newly synthesized RNA population comprising the labeled nucleotide analog;

hybridizing a probe and the newly synthesized RNA population comprising the labeled nucleotide analog, under stringent hybridization conditions, wherein the probe is complementary to at least a portion of a transcription initiation region of the newly synthesized RNA population;

digesting unhybridized, single-stranded RNA with a single-strand specific ribonuclease to form a protected RNA population;

detecting a protected RNA population comprising the labeled nucleotide analog; and

quantitating the protected RNA population comprising the labeled nucleotide analog.

[0034] Also provided is a method for determining whether a test compound is an RNA synthesis initiation inhibitor of a positive strand RNA virus comprising:

contacting an isolated replicase complex for the positive strand RNA virus comprising a viral replicon template RNA for the positive strand RNA virus, a labeled nucleotide analog, and the test compound, under conditions sufficient for *in vitro* RNA synthesis, to form a newly synthesized RNA population comprising the labeled nucleotide analog;

hybridizing a probe and the newly synthesized RNA population comprising the labeled nucleotide analog, under stringent hybridization conditions, wherein the probe is complementary

to at least a portion of an initiation region of the newly synthesized RNA population;
digesting unhybridized, single-stranded RNA with a single-strand specific ribonuclease to form a protected RNA population;
detecting a protected RNA population comprising the labeled nucleotide analog;
quantitating the protected RNA population comprising the labeled nucleotide analog to provide a test RNA amount; and
comparing the test RNA amount with a control RNA amount of protected RNA comprising the labeled nucleotide analog but produced in the absence of the test compound, wherein a decrease in the test RNA amount compared to the control RNA amount indicates that the test compound inhibits RNA synthesis initiation of the positive strand RNA virus.

[0035] In some embodiments of the above methods, the isolated viral replicase complexes comprising a viral replicon template RNA may be provided by transfecting a cell line with an isolated DNA template for a viral replicon or an isolated viral RNA, to provide a transfected cell line, incubating the transfected cell line under conditions suitable for viral replication, and isolating the replicase complexes comprising viral replicon template RNA from the cell membrane fraction of the transfected cell line.

[0036] In other embodiments of the above methods, the isolated viral replicase complexes comprising viral replicon template RNAs may be provided by acutely infected or persistent infected primary hepatocytes, lymphocytes or other cell lines, incubating the infected cell line under conditions suitable for viral replication, and isolating the replicase complexes comprising viral replicon RNA from the cell membrane fraction. The replication complexes may be isolated from infected primary cells or cell lines.

[0037] In certain embodiments of the methods described herein, the positive strand RNA virus is Hepatitis C Virus, and a suitable DNA template for the HCV viral replicon is, for example, SEQ ID NO: 1, which can also be described as GenBank Accession No. AJ242652. Other suitable replicon template DNAs include, for example, AB114136, AJ242654, AJ242653, and AJ242651 (SEQ ID NOs. 2-5).

[0038] In certain embodiments of the methods, the labeled nucleotide analog is an analog capable of being recognized by a specific antibody, an analog which can be recognized via a high specificity binding reaction, or an analog directly detectable as a result of a physical property of the analog. In some embodiments, the labeled nucleotide analog is an analog capable of being recognized by a specific antibody, such as 5'-bromouridine 5'-triphosphate (Br-UTP).

In other embodiments, the labeled nucleotide analog is a radioactively labeled nucleotide.

[0039] In the methods described herein, detection of a newly synthesized RNA population may be accomplished by contacting a newly synthesized RNA population with an antibody specific for the labeled nucleotide analog, and immuno-precipitating the newly synthesized RNA population comprising the labeled nucleotide analog to form an immuno-precipitated RNA population. In some embodiments, the nucleotide analog is Br-UTP, and the specific antibody is an anti-BrdU antibody. The methods described herein include quantitating the newly synthesized RNA population by performing real time PCR on the immuno-precipitated RNA population. In other embodiments, detection of a newly synthesized RNA population may be accomplished by detecting a radioactively labeled nucleotide by, for example, autoradiography.

[0040] In some embodiments, a 2'-O-methylated nucleotide may be employed during the replication reaction to increase the yield of newly synthesized RNA.

[0041] In order to form a newly synthesized RNA population, an isolated replicase complex comprising a viral replicase RNA template is employed. The isolated viral replicon template RNA comprises a transcription initiation region. The template RNA may be, for example, isolated as a component of the replicase complex. In some embodiments, an isolated viral template RNA may be added to the replicase complex.

[0042] A viral replicon RNA or a DNA vector encoding the viral replicon RNA may be transfected into a cell line suitable for expression of the viral replicon RNA. Suitable cell lines include, for example, mammalian cell lines such as Vero cells, HeLa cells, CHO cells, COS cells, WI38 cells, N1H-3T3 cells (and other fibroblast cells, such as MRC-5 cells), MDCK cells, KB cells, SW-13 cells, MCF7 cells, BHK cells, HEK-293 cells, HepG2 cells, Bowes melanoma cells cell lines; and chicken embryonic fibroblast (CEF) cell lines. Preferred cell lines include human cell lines, such as, for example, human hepatosoma cell lines such as Huh-7. Suitable means of transfection include, for example, calcium phosphate co-precipitates, mechanical procedures such as microinjection, electroporation, insertion of a plasmid encased in liposomes, as well as other techniques known in the art.

[0043] A suitable DNA vector for production of an HCV replicon RNA, for example, comprises, 5'-3', the HCV-IRES, the neomycin phosphotransferase (neo) gene, the IRES of encephalomyocarditis virus, which directs translation of HCV sequences NS3 to NS5B, and the 3'-NTR. The sequence of a DNA vector suitable for production of the HCV replicon RNA has

been deposited in GenBank (Accession no. AJ242652) (SEQ ID NO:1). The HCV replicon DNA vector may be transfected into Huh-7 cells by electroporation. HCV replicon RNA may be transcribed *in vitro* using either SP6 RNA polymerase or T7 RNA polymerase from a plasmid DNA comprising the appropriate promoter with the cDNA encoding the replicon RNA downstream of the promoter. The DNA template for production of a viral replicase RNA may be produced by PCR amplification using a DNA template encoding the viral replicon and specific primers one of which comprises an SP6 or T7 RNA polymerase promoter. Alternatively, the DNA template for production of a viral replicase RNA may be produced by reverse transcription of the viral replicase RNA using primers one of which comprises an SP6 or T7 RNA polymerase promoter.

[0044] Once the viral replicon RNA or DNA vector encoding the viral replicon RNA is transfected into a cell, the cell culture may be grown to confluency. The cells in the culture produce replicase complexes comprising viral replicon template RNAs, which can then be isolated from the cells. Viral replicon RNA includes minus strand viral replicon template RNA, and optionally positive strand RNA.

[0045] Alternatively, infected cells such as acutely infected or persistent infected primary hepatocytes, lymphocytes or other cell lines. Infected primary cells or cell lines may be employed. The infected cell line is incubated under conditions suitable for viral replication, and the replicase complexes comprising viral replicon RNA are isolated from the cell membrane fraction.

[0046] In order to isolate the replication complexes comprising viral replicon template RNA, the cells (i.e., transfected or infected) may be lysed and centrifuged at low speed (e.g., about 900 X to about 1000 X g) to remove cellular debris. The supernatant may then be centrifuged at higher speed (e.g., about 15000 X g) to obtain the membrane fraction, which comprises isolated viral replicon complexes comprising viral replicon template RNA. It has been shown that replication of the HCV RNA occurs in a membrane-bound replication complex. Thus, HCV replicon complexes and replicon template RNA can be isolated by isolating the membrane fraction of the cells expressing the HCV replicon RNA.

[0047] The isolated membrane fraction containing replicase complexes comprising viral replicon template RNAs (e.g., an aliquot of the membrane fraction of a cell line expressing a viral RNA replicon) is used to perform RNA replication *in vitro* in a cell-free system. The isolated replicase complexes comprising viral replicon template RNA are contacted with a

labeled nucleotide analog under conditions sufficient to promote *in vitro* RNA synthesis. Conditions sufficient to promote *in vitro* RNA synthesis include suitable buffers, and nucleoside triphosphates (i.e., ATP, UTP, CTP, and GTP). The labeled nucleotide analog is incorporated into a newly synthesized RNA population, and provides a means for detecting only the newly synthesized RNA, and not the template RNA. In one embodiment, suitable conditions for replication include, for example, incubation in 50 μ l (total volume) of 50 mM HEPES (pH 7.5), 10 mM KCl, 10 mM MgCl₂, 50 units RNasin, 10 μ g/ml actinomycin D, 2.5 mM ATP, 0.5 mM each CTP, GTP, labeled UTP and an aliquot of a membrane fraction of a cell line expressing the viral replicon RNA, e.g., the HCV replicon. In another embodiment, suitable replication conditions include 50 mM HEPES (pH 7.3); 10 mM KCl; 10 mM MgCl₂; 0.3 mM MnCl₂; 20 Units of RNase inhibitor; 10 μ g of actinomycin D per mL; 0.5 mM ATP, GTP, and UTP; 10 μ Ci of [α -P³²] CTP; and 6 μ l of the membrane fraction of a cell line expressing the viral replicon RNA in a total volume of 60 μ l. The replication reaction is continued for a time sufficient to produce a desired amount of newly synthesized RNA.

[0048] In some embodiments, the replication conditions may include adding a 2'-O-methylated nucleotide such as, for example, 2'-O-methyl-5-methyluridine-5'-triphosphate to the replication mixture. The inclusion of a 2'-O-methylated nucleotide such as 2'-O-methyl-5-methyluridine-5'-triphosphate may increase the efficiency of the replication complex, resulting in an increase in the amount of labeled product obtained in the absence of the 2'-O-methyl-5-methyluridine-5'-triphosphate. Without being held to a particular theory, it is believed that a 2'-O-methylated nucleotide, particularly 2'-O-methyl-5-methyluridine-5'-triphosphate increases the efficiency of the replicase complex when a radiolabeled nucleotide such as, for example, P³²-CTP is added to the reaction at low concentrations.

[0049] Suitable labeled nucleotide analogs include, for example, analogs which are capable of being recognized by specific antibodies (e.g., Br-UTP), analogs such as a biotin labeled nucleotide (e.g., biotin-CTP) which can be recognized via a high specificity binding reaction (i.e., biotin/avidin or biotin/streptavidin binding), and analogs which are directly detectable as a result of a physical property of the analog, such as radioactivity (e.g., a P³²-labeled nucleotide), fluorescence, luminescence etc., for example a fluorescein-nucleotide (e.g., fluorescein-UTP). In one embodiment, the labeled nucleotide analog is 5'-bromouridine 5'-triphosphate (Br-UTP). The RNA containing Br-UTP may be isolated by immuno-precipitation and/or detection with an anti-BrdU antibody. The anti-BrdU antibody may optionally be labeled

with a tag suitable for direct detection of the antibody such as, for example, fluorescein, or other dyes such as those available from Molecular Probes. In other cases, the antibody may be detected via binding to a labeled secondary antibody such as, for example, a rabbit or mouse IgG labeled with alkaline phosphatase or horseradish peroxidase.

[0050] The labeled nucleotide analog is added to the replication mixture so that the labeled nucleotide analog is incorporated into the newly synthesized RNA (e.g., newly synthesized viral replicon RNA) in place of at least a portion of the corresponding unlabeled nucleotide. For example, Br-UTP may be used in place of at least a portion of the UTP. In certain cases, all of the corresponding nucleotide may be replaced with the labeled nucleotide analog. After replication, there are two populations of RNA (e.g., viral replicon RNA): an unlabeled RNA population and a labeled RNA population. The newly synthesized RNA population, which contains the labeled nucleotide analog, can be distinguished from the RNA population which was present before the replication step, because the RNA population which was present before replication will not contain the labeled nucleotide analog. Because the newly synthesized RNA can be distinguished from the RNA present prior to replication, the amount of newly synthesized RNA (e.g., viral replicon RNA) can be detected and/or quantified by employing a means of RNA detection and/or quantitation. The means for RNA detection and quantitation may be the same or different.

[0051] After replication, the RNA in the replication mixture optionally may be purified using, for example, a commercially available RNA purification kit. The newly synthesized RNA is then detected using a means for detecting the RNA. Detection preferably comprises quantitative detection. The method for detecting the RNA is selected based on the labeled nucleotide analog employed. For example, in the case of an analog capable of being recognized by specific antibodies, a specific antibody may be employed to immuno-precipitate the labeled RNA. The immuno-precipitated RNA may be detected, for example, by directly detecting a fluorescent tag present on the antibody. Alternatively, the immuno-precipitated RNA may be detected with a labeled secondary antibody. In yet another alternative, the immuno-precipitated labeled RNA can be amplified and detected using real-time PCR.

[0052] For example, a newly synthesized RNA population containing Br-UTP as the labeled nucleotide analog may be immuno-precipitated using an anti-BrdU monoclonal antibody. The unlabeled, template RNA will not be precipitated, while the newly synthesized Br-UTP-labeled RNA will be specifically precipitated. The immuno-precipitated, labeled RNA may then

be quantified using real time PCR. Thus, in the case of, for example, a Br-UTP labeled viral replicon RNA precipitated with anti-BrdU antibody, the method for detection comprises contacting the population of viral replicon RNAs with an anti-BrdU antibody, precipitating the viral replicon RNAs, and detecting and quantitating the labeled viral replicon RNAs using a suitable method such as, for example, real time PCR.

[0053] Real time PCR (polymerase chain reaction) is a quantitative reverse transcription-PCR reaction (RT-PCR). Early in RT-PCR, reagents are in excess, template and product are at low enough concentrations that product renaturation does not compete with primer binding, and amplification proceeds at a substantially constant, exponential rate. At some variable point during the reaction, the reaction rate ceases to be exponential, and enters a linear phase of amplification. Late in the amplification cycle, little product is made. Real time PCR allows the collection of data during the exponential phase of amplification, and thus allows for accurate quantitation of the amount of amplified product. Detection and quantitation of the double-stranded DNA produced in the PCR reaction is done using a fluorescent reporter, the signal of which increases in direct proportion to the amount of double-stranded DNA PCR product in a reaction. Taqman® probes labeled with fluorescent dye such as FAM (6-carboxy-fluorescein) may be used. Another suitable reporter is the double-strand DNA-specific dye SYBR® Green (Molecular Probes). SYBR® Green binds double-stranded DNA, and upon excitation emits light. Thus, as a PCR product accumulates, fluorescence increases. Other alternative reporters include hybridization probes which use fluorescence energy transfer (FRET) for detection.

[0054] When the labeled nucleotide analog is a radioactive nucleotide such as P³²-CTP, the labeled RNAs may be detected by autoradiography, direct radioactive detection such as Phosphorimaging, and the like.

[0055] In the case of a labeled nucleotide analog which can be recognized via a high specificity binding reaction, the labeled RNAs may be isolated using the high specificity binding reaction. For example, a biotin-labeled nucleotide can be recognized by avidin or streptavidin. Once the labeled RNAs have been isolated, the labeled RNAs can be amplified and detected using real-time PCR. In the case of a directly detectable nucleotide analog, the labeled nucleotide analog can be directly detected using fluorescence, luminescence, and the like.

[0056] Suitable controls for quantitation of the newly synthesized RNA population follow. Template RNAs isolated from cells untransfected with a DNA vector or viral RNA and subjected to replication, have no detectable amount of newly synthesized RNA. Template RNAs

isolated from cells transfected with a template DNA vector or viral RNA, but replicated without the labeled nucleotide analog, have only background levels of newly synthesized RNA.

Template RNAs from cells transfected with template DNA vector or viral RNA, replicated in the presence of a labeled nucleotide analog, which are not detected with the means for detection, show only background levels of newly synthesized RNA. Template RNAs from cells transfected with template DNA vector or viral RNA, replicated in the presence of labeled nucleotide analog and detected with the means for detection should, however, show a strong newly synthesized RNA signal in this assay.

[0057] The assay may be employed to quantitate a newly initiated RNA population. In the case of HCV, for example, most (e.g., about 90%) of the newly synthesized viral RNAs are the elongation products of previously initiated (i.e., initiated) template RNAs. In an initiated template RNA, a portion of the newly synthesized RNA is previously made using a minus strand template. Upon addition of the appropriate factors (e.g., buffers and nucleoside triphosphates), elongation of the initiated RNA can proceed to form the full-length RNA. Only a small fraction (e.g., less than about 10%) of the newly synthesized RNAs represent newly initiated products formed from uninitiated template RNAs (i.e., pre-initiation). Both the previously initiated and newly initiated RNAs will comprise the labeled nucleotide analog in the elongation region of the RNA. Only the newly initiated RNAs, however, will comprise the labeled nucleotide analog in the transcription initiation region. A method to select and quantitate newly initiated RNAs comprises employing RNA protection to the newly synthesized RNA population which may comprise previously initiated RNA population as well as newly initiated RNAs.

[0058] In the RNase protection assay, a nucleic acid probe is added to the newly synthesized viral replicon RNA population which may comprise previously initiated RNAs as well as newly initiated RNAs. The nucleic acid probe comprises a region complementary at least a portion of the transcription initiation region of the viral replicon RNAs (i.e., the newly synthesized RNA population). When the probe is hybridized to the population of newly synthesized viral replicon RNAs, a portion of the transcription initiation region becomes double-stranded, with the remainder of the RNA being single-stranded. The single-stranded RNA may then be removed, for example, by digestion with a single-strand specific ribonuclease specific for single-stranded RNA. Suitable single-strand specific ribonucleases include, for example, ribonuclease T1, ribonuclease A, nuclease S1, and combinations comprising one or more of the foregoing single-strand specific ribonucleases. After digestion, the remaining RNA will thus be

a double-stranded RNA (i.e., protected) population in which the previously initiated RNAs will comprise no labeled nucleotide analog, and the newly initiated RNAs comprise the labeled nucleotide analog. The protected newly initiated RNAs can then be detected and/or quantified as described above.

[0059] Hybridization of the probe may be performed under conditions which are sufficient to allow specific hybridization of the nucleic acid probe to the viral replicase RNA. Specific hybridization can be performed under high stringency conditions or moderate stringency conditions, for example.

[0060] In a particularly preferred embodiment, the hybridization conditions for specific hybridization are high stringency.

[0061] The assay described herein may be employed to screen test compounds for inhibition of viral replicon RNA synthesis. Such compounds include replicase complex activity inhibitors as well as RNA synthesis initiation and elongation inhibitors. The method may be employed to distinguish initiation inhibitors from elongation inhibitors. If a viral replicon RNA synthesis inhibitor is added to a replication mixture, less viral replicon RNA should be synthesized, and the amount of labeled viral replicon RNA detected will be decreased (e.g., test RNA amount) relative to the amount of viral replicon RNA in a control sample (i.e., control RNA amount) with no added inhibitor. The amount of labeled viral replicon RNA detected will be decreased, for example, if the viral replicon RNA synthesis inhibitor is a viral replicon RNA elongation inhibitor. Viral replicon RNA elongation inhibitors will block synthesis of newly initiated viral replicon RNA elongation products and also previously initiated viral replicon RNA elongation products. If, however, the viral replicon RNA synthesis inhibitor is an initiation inhibitor, signal strength will be only modestly affected. Most (e.g., approximately 90%) of the newly synthesized viral RNAs of HCV, for example, are the elongation products of previously initiated viral replicon RNAs. Thus, only a small fraction (e.g., less than about 10%) of the newly synthesized replicon RNAs will be affected by an initiation inhibitor. FIGURE 1 provides a schematic of the inhibition of HCV RNA synthesis by initiation and elongation inhibitors.

[0062] An RNase protection assay, as described herein, can be added after formation of the newly synthesized RNA population to identify initiation inhibitors. In this assay, a nucleic acid probe which is complementary to at least a portion of the initiation region of the newly synthesized viral replicon is employed. The probe is hybridized to the viral replicon RNA such that the region of the viral replicon RNA complementary to the probe is double-stranded (i.e.,

replicated), and the remainder of the viral replicon RNA is single-stranded. Hybridization is preferably performed under stringent conditions. The single-stranded portion of the viral replicon RNA is then digested with a ribonuclease specific for single-stranded RNA such as, for example, ribonuclease A, ribonuclease T1, nuclease S1, or a combination comprising one or more of the foregoing ribonucleases. The remaining population of double-stranded replicon RNA will then contain two fractions: a fraction of previously initiated viral replicon RNAs which comprise no labeled nucleotide analog, and a fraction of viral replicon RNAs initiated after addition of the labeled nucleotide analog which comprise the labeled nucleotide analog. The protected viral replicon RNAs containing the labeled nucleotide analog can then be specifically detected and/or quantified as described above.

[0063] In the case of HCV, an exemplary probe is an RNA complimentary to nucleotides 15 to 434 of the HCV replicon RNA (SEQ ID NO:6), which hybridizes with the newly synthesized viral replicase RNA near the initiation codon. The RNA not hybridized to the probe is digested with a single-stranded ribonuclease, leaving only nucleotides 15 to 434 of the replicon RNA. The digested RNA will have two double-stranded RNA populations: a population of previously initiated replicon RNAs which do not contain the labeled nucleotide analog, and a population of newly initiated replicon RNAs which contain the labeled nucleotide analog. Thus, previously initiated HCV RNA elongation products will not be detected, while newly synthesized RNAs will be detected. In this assay including RNA protection, when an initiation inhibitor is added to the replication mixture, a reduced amount of labeled nucleotide analog is incorporated at the initiation site, and the signal measured by the means for detection will be decrease decreased.

[0064] In positive strand RNA viruses, the replicase complex is a multi-protein complex which replicates viral replicon RNA from a viral replicon template RNA. Replicase complexes which include viral replicon template RNAs in addition to replication proteins can be isolated as a membrane bound fraction from cells expressing the viral replicon RNA. The isolated replicon complexes include both previously initiated and uninitiated viral replicon template RNAs. Upon addition of nucleotide triphosphates and other components sufficient for RNA synthesis, elongation of previously initiated template RNAs proceeds to produce full-length viral replicase RNAs. It had not, however, been demonstrated that *de novo* initiation of uninitiated viral replicon template RNAs occurs in isolated replicase complexes. By using an inhibitor of initiation of viral replicon RNA synthesis initiation, it can be shown that *de novo* initiation of

viral replicase RNA synthesis occurs in isolated replicase complexes. Because *de novo* initiation of RNA synthesis takes place in the isolated replicase complexes, the methods may be used to quantify newly initiated RNA and also to identify inhibitors of RNA synthesis initiation.

[0065] Thus, the described methods are suitable for detecting inhibitors of RNA synthesis of positive strand RNA viruses such as RNA initiation and elongation inhibitors.

[0066] Also provided herein is a kit for screening a test compound for inhibition of RNA synthesis of a positive strand RNA virus. The kit comprises a replicase complex for the positive strand RNA virus comprising a viral replicon template RNA for the positive strand RNA virus, instructions for use of the kit, and buffers and nucleoside triphosphates which are sufficient for synthesis of the viral replicon RNA. Instructions may include, for example, instructions for quantitating a newly synthesized RNA population, instructions for determining if a compound is an RNA synthesis inhibitor, and the like. The instructions may be written instructions, although electronic storage media (e.g., magnetic diskette or optical disk) containing instructions are also acceptable, relating to the use of components of the methods. The instructions included with the kit may include information as to reagents (e.g., whether included or not in the kit) necessary for practicing the methods, instructions on how to use the kit, and/or appropriate reaction conditions.

[0067] Suitable buffers include those described previously as suitable for performing replication with isolated replication complexes.

[0068] The kits may optionally further comprise, for example, a labeled nucleotide analog, a means for detecting and/or quantitating a labeled nucleotide analog, a primer, a single-strand specific ribonuclease, and instructions for performing an RNase protection assay.

[0069] The component(s) of the kit may be packaged in a convenient, appropriate packaging. The components may be packaged separately, or in one or multiple combinations.

[0070] The invention is further illustrated by the following nonlimiting examples.

EXAMPLES

EXAMPLE 1. GROWTH AND MAINTENANCE OF HCV REPLICON CONTAINING CELLS *1A. HCV Replicon Expression and Cell Maintenance*

[0071] RNA molecules encoding the HCV replicon (i.e., viral replicon RNAs) are transfected into Huh-7 cells using electroporation.

[0072] The equipment and materials for cell maintenance include, but are not limited to, Huh-7 HCV replicon-containing cells, maintenance media (DMEM (Dulbecco's modified Eagle

media) supplemented with 10% FBS, L-glutamine, non-essential amino acids, penicillin (100 units/ml), streptomycin (100 micrograms/ml), and 500 micrograms/ml of Geneticin (G418), screening media (DMEM supplemented with 10% FBS, L-glutamine, and non-essential amino acid, penicillin (100 units/ml) and streptomycin (100 micrograms/ml)), 96 well tissue culture plates (flat bottom), 96 well plates (U bottom for drug dilution), Interferon alpha for positive control, fixation reagent (such as methanol: acetone), primary antibody (rabbit anti-NPTII), secondary antibody: Eu-N1 1, and enhancement solution.

[0073] Cells containing the HCV replicon support high levels of viral RNA replication when their density is suitable. Over-confluency may cause decreased viral RNA replication. Therefore, cells should be grown in log phase in the presence of 500 micrograms/ml of G418. Generally, cells should be passed twice a week at 1: 4-6 dilution. Cell maintenance is conducted as follows:

[0074] Cells containing the HCV replicon DNA vector are examined under a microscope to ensure that the cells are growing well. Cells are rinsed once with phosphate buffered saline (PBS) and 2 ml trypsin is added. The cell/ trypsin mixture is incubated at 37°C in a CO₂ incubator for 3-5 minutes. After incubation 10 ml of complete media is added to stop the trypsinization reaction. Cells are blown gently, put into a 15 ml tube, and spun at 1200 rpm for 4 minutes. The trypsin/ medium solution is removed and the pelleted cells are recovered.

EXAMPLE 2. QUANTITATIVE ASSAY FOR DETECTION OF NEWLY SYNTHESIZED HCV RNA IN CELL-FREE SYSTEM

[0075] Membrane Fractions are purified from Huh-7 cells transfected with the HCV replicon RNA according to the procedure given by Hardy, et al. (*J. Virol.* (2003) 77:2029-2037). Briefly, the HCV replicon-containing cells are washed with 1X PBS, re-suspended in cold hypotonic buffer (10mM Tris-HCl, pH 7.8, 10 mM NaCl), and put on ice for 20 minutes. The swelled cells are disrupted using a dounce homogenizer. The mix is centrifuged at 900 X g for 5 minutes at 4°C. The supernatant is transferred to a fresh tube and centrifuged at 15000 X g for 25 minutes at 4°C. The pellet, which contains the membrane fraction, is re-suspended in storage buffer (hypotonic buffer with 15% glycerol), and may be stored at -80°C.

[0076] *In vitro* RNA replication is performed according to the procedure of Lai (*J. Virol.* (2003) 77:2295-2300) with some modifications. Test compound (1 μM to 100 μM in 0.5 μl

DMSO) is pre-incubated the membrane fraction (10 μ l) for 10 minutes at 30°C before the replication mix is added. The replication mix (50 μ l total volume) contains 50 mM HEPES (pH 7.5), 10 mM KCl, 10 mM MgCl₂, 50 U RNasin, 10 μ g/ml actinomycin D, 2.5 mM ATP, 0.5 mM each CTP, GTP, Br-UTP and the pre-incubated membrane fraction (i.e., the replicase complex and viral replicon template RNA). The replication mix is incubated at 30°C for 2 hours, followed by purification using TRIZOL (Invitrogen, Carlsbad, CA) or RNA EASY (Qiagen, Valencia, CA) according to manufacturer's instructions. The RNA samples are then immuno-precipitated using anti-BrdU monoclonal antibody (Molecular Probes, Eugene, OR) and protein A agarose (Invitrogen) in 200 μ l to 400 μ l of precipitation buffer containing 1X PBS, 0.05% NP 40, 0.1 μ g/ μ l tRNA, 5 U/ μ l RNasin, and 0.3 M uridine.

[0077] When the newly synthesized segment near the initiation site is of interest, an RNA protection assay is used after Br-UTP labeling and RNA purification. The RNA probe for segment from 15 to 434 nt (SEQ ID NO: 6) is synthesized using an *in vitro* transcription kit (MAXIscript, Ambion). The DNA template for *in vitro* transcription was made by PCR using forward primer 5' GGGGGCGACACTCCACCATAGAT (15-37) (SEQ ID NO: 7) and reverse primer 5' ATTTAGGTGACACTATAGAAACCCAAGCGGCCGGAGAACCT (413-434, plus SP6 core sequence) (SEQ ID NO:8) on DNA vector for replicon HCV RNA or cDNA made from replicon HCV RNA by reverse transcription. The RNA is protected using an RNA protect assay kit (e.g. RPA III, Ambion, Austin, TX). Briefly, in the RNA protection assay, the newly synthesized replicon RNAs are hybridized to the RNA probe. The hybridized RNAs are then treated with a mixture of ribonuclease A and ribonuclease T1 to digest the single-stranded RNA. The remaining double-stranded RNA may then be precipitated prior to detecting.

[0078] The newly synthesized HCV RNA produced in the cell-free system is quantitatively detected using real time PCR. Table 1 shows the data for several control experiments. A Taqman® probe labeled with fluorescent dye FAM was used.

TABLE 1		
Sample	Copy number after PCR with no immuno-precipitation with anti-BrdUTP antibody	Copy number after PCR performed after immuno-precipitation with anti-BrdUTP antibody
Template RNA isolated from Huh-7 cells expressing HCV replicon with Br-UTP	1.86×10^7	2.05×10^5
Template RNA isolated from Huh-7 cells expressing HCV replicon with Br-UTP: No anti-Brd-UTP antibody	1.85×10^7	2401
Membranes isolated from Huh-7 cells with no HCV (no HCV template RNA)	15	3
Template RNA isolated from Huh-7 cells expressing HCV replicon: No Br-UTP in replication	1.79×10^7	7270

[0079] As seen from Table 1, for template RNA isolated from Huh-7 cells expressing the HCV replicon with Br-UTP in the replication mixture, a high copy number of replicon RNA is observed both with and without precipitation with an anti-BrdUTP antibody. If no anti-BrdUTP antibody is used, the sample without precipitation has a high copy number of HCV replicon RNA, while the precipitated population has only a background amount. If the Huh-7 cells are not transfected with the HCV replicon (i.e., no template RNA), a background amount of replicon RNA is observed with or without precipitation with an anti-BrdU antibody. If there is no Br-UTP in the replication mixture with template RNA isolated from Huh-7 cells expressing the HCV replicon, the sample without precipitation has a high copy number of HCV replicon RNA, while the precipitated population has only a background amount. Thus, the control experiments give the expected results.

[0080] As shown in Table 2, the copy number of the HCV RNA measured by quantitative PCR is dependent upon the amount of template RNA added to the replication mixture. In Table 2, HCV is a replication mixture with template RNA isolated from Huh-7 cells expressing the HCV replicon. HVC (1/2) and HCV (1/8) are replication mixtures containing $\frac{1}{2}$ and $\frac{1}{8}$ of the amount of template RNA, respectively. Huh-7 (no HCV) is a no HCV template control as described above. No Br-UTP is a control in which template RNA isolated from Huh-7

cells expressing the HCV replicon, but no Br-UTP is added to the replication mixture.

TABLE 2		
Sample	Copy number after PCR with no immuno-precipitation with anti-BrdUTP antibody	Copy number after PCR performed after immuno-precipitation with anti-BrdUTP antibody
HCV	2.19×10^7	7.34×10^5
HCV (1/2)	1×10^7	3.7×10^5
HCV (1/8)	3.1×10^6	1.51×10^5
Huh-7 (No HCV)	19	2030
No Br-UTP	2.17×10^7	4.6×10^4

[0081] As shown in Table 2, the copy number of HCV replicons measured in the assay, both with and without precipitation with an anti-BrdUTP antibody, is directly proportional to the amount of template replicon RNA added to the reaction mixture.

[0082] Table 3 shows the data for a known HCV RNA replication inhibitor that is tested using the disclosed assay at a concentration of 10 μ M. ACH-0134152 is a published compound (Dhanak et al., *J. Biol. Chem.* (2002) 41):38322-7. The structure of ACH-0134152 is shown below.

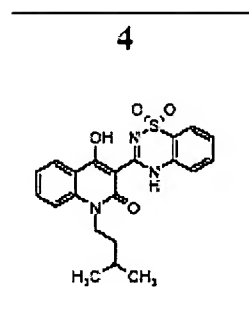


TABLE 3				
Sample	Copy number after PCR with no RNase protection assay		Copy number after PCR with RNase protection assay	
		Copies as % of samples with no test compound		Copies as % of samples with no test compound
No test compound	46942	100	137494	100
ACH-0134152	31156	66	16970	12

[0083] In Table 3, although treatment with ACH-0134152 results in less than half reduction in total viral RNA synthesis with no RNase protection, it causes close to a 10-fold reduction in the amount of newly initiated viral RNA synthesis after performing RNase protection. Thus, compound ACH-0134152 inhibits HCV RNA synthesis at the initiation stage.

[0084] The results reported in this example validate this assay for compound screening and profiling. The assay has the advantage of quantitatively detecting newly synthesized HCV RNA, and permitting but also specific segment of our choice.

EXAMPLE 3. USE OF AN RNA INITIATION INHIBITOR TO DEMONSTRATE *DE NOVO* INITIATION OF REPLICATION IN ISOLATED REPLICATION COMPLEXES

[0085] An RNA replication assay was used to determine if the replication complexes isolated from a membrane fraction were enzymatically active. The ACH-0134152 initiation inhibitor was employed to determine if the isolated replication complexes were capable of *de novo* initiation of newly synthesized RNA. Standard replication mixtures contained 50 mM HEPES (pH 7.3); 10 mM KCl; 10 mM MgCl₂; 0.3 mM MnCl₂; 20 Units of RNase inhibitor; 10 µg of actinomycin D per mL; 0.5 mM ATP, GTP, and UTP; 10 µCi of [α -P³²] CTP; and 6 µl of the membrane fraction in a total volume of 60 µl. The reaction mixtures were incubated at 30°C for 2 hours. The RNA products were extracted with phenol-chloroform, ethanol precipitated, and separated on a 1% agarose gel. After electrophoresis, the gel was fixed with 10% glacial acetic acid then in ethanol and dried prior to autoradiography.

[0086] Figure 2 shows an autoradiogram of the experimental results. Lane 4 is a control lane with no added inhibitor. Lanes 1-3 show the results with initiation inhibitor concentrations of 100, 20, and 4 µM, respectively. DS stands for double-stranded RNA and SS stands for single-stranded RNA. As clearly shown in the autoradiogram, addition of the initiation inhibitor reduces the amount of single-stranded RNA produced by the replication complex. In the assay, double-stranded RNA is produced largely from elongation of previously initiated RNAs, while single-stranded RNA is produced from newly initiated RNAs. The presence of a single-stranded RNA band in control lane 4 shows that newly initiated RNAs are formed during replication. When ACH-0134152, a known RNA synthesis initiation inhibitor, is added during replication, the single-stranded RNA band disappears, while the double-stranded RNA band remains. Thus, by employing an RNA synthesis initiation inhibitor, it is confirmed that *de novo* RNA synthesis

initiation occurs in isolated replicase complexes. .

[0087] A method has been described to identify inhibitors of viral replicon RNA synthesis using detection and/or quantitation of a population of newly synthesized RNA. A decrease in the amount of newly synthesized RNA made in the presence of a test compound compared to a control with no test compound indicates that the test compound is a viral replicon RNA synthesis inhibitor. The method is particularly useful for identifying RNA elongation inhibitors. In addition, an RNase protection assay may be employed on the newly synthesized RNA to identify RNA initiation inhibitors. It has been clearly demonstrated that isolated replication complexes are capable of *de novo* initiation in addition to elongation of previously initiated RNA. An advantage of the method is that initiation inhibitors can be distinguished from elongation inhibitors. Another advantage is that quantitation of the newly synthesized RNA population is employed to identify viral replicon RNA synthesis inhibitors. Another advantage is that a 2'-O-methylated nucleotide can be used to increase the yield of newly synthesized RNA produced during replication.

[0088] From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention.

CLAIMS

What is claimed is:

1. A method for determining whether a test compound inhibits RNA synthesis of a positive strand RNA virus, comprising:
 - contacting an isolated replicase complex for the positive strand RNA virus comprising a viral replicon template RNA for the positive strand RNA virus, a labeled nucleotide analog, and the test compound, under conditions sufficient for *in vitro* RNA synthesis, to form a newly synthesized RNA population comprising the labeled nucleotide analog;
 - detecting the newly synthesized RNA population comprising the labeled nucleotide analog;
 - quantitating the newly synthesized RNA population comprising the labeled nucleotide analog to provide a test RNA amount; and
 - comparing the test RNA amount with a control RNA amount of a newly synthesized RNA population comprising the labeled nucleotide analog produced in the absence of the test compound, wherein a decrease in the test RNA amount compared to the control RNA amount indicates that the test compound inhibits RNA synthesis of the positive strand RNA virus.
2. The method of Claim 1, wherein contacting further comprises contacting with 2'-O-methyl-5-methyluridine-5'- triphosphate.
3. The method of Claim 1, further comprising providing the isolated replicase complex comprising the viral replicon template RNA by:
 - transfecting a cell line with a viral replicon RNA or a DNA template for a viral replicon to provide a transfected cell line,
 - incubating the transfected cell line under conditions suitable for production of viral replicase complexes, and
 - isolating the replicase complexes and the viral replicon template RNA from the cell membrane fraction of the transfected cells.

4. The method of Claim 3, wherein the positive strand RNA virus is Hepatitis C Virus and the DNA template for a viral replicon is SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO:3, SEQ ID NO: 4, or SEQ ID NO: 5.

5. The method of Claim 1, further comprising providing the isolated replicase complex comprising the viral replicon template RNA by
incubating a positive strand RNA virus infected primary cell or cell line under conditions suitable for production of viral replicase complexes, and
isolating the replicase complexes comprising the viral replicon template RNA from the cell membrane fraction of the infected cells.

6. The method of Claim 1, wherein the positive strand RNA virus is Hepatitis C Virus.

7. The method of Claim 1, wherein the labeled nucleotide analog is an analog capable of being recognized by a specific antibody, an analog which can be recognized via a high specificity binding reaction, or an analog directly detectable as a result of a physical property of the analog.

8. The method of Claim 7, wherein the labeled nucleotide analog is an analog capable of being recognized by a specific antibody.

9. The method of Claim 8, wherein the labeled nucleotide analog is Br-UTP.

10. The method of Claim 8, wherein detecting comprises contacting the newly synthesized RNA with an antibody specific for the labeled nucleotide analog and immuno-precipitating the newly synthesized RNA comprising the labeled nucleotide analog to form immuno-precipitated RNA.

11. The method of Claim 10, wherein the labeled nucleotide analog is Br-UTP and the specific antibody is an anti-BrdU antibody.

12. The method of Claim 10, wherein quantitating the newly synthesized RNA comprises performing real time PCR on the immuno-precipitated RNA.
13. The method of Claim 1, wherein the test compound is an RNA elongation inhibitor.
14. The method of Claim 1, wherein the test compound is an RNA synthesis initiation inhibitor.
15. The method of Claim 1, wherein the test compound is a replicase complex activity inhibitor.
16. A method for quantitating newly initiated RNA of a positive strand RNA virus comprising:
 - contacting an isolated replicase complex for the positive strand RNA virus comprising a viral replicon template RNA for the positive strand RNA virus, and a labeled nucleotide analog, under conditions sufficient for *in vitro* RNA synthesis, to form a newly synthesized RNA population comprising the labeled nucleotide analog;
 - hybridizing a probe and the newly synthesized RNA population comprising the labeled nucleotide analog, under stringent hybridization conditions, wherein the probe is complementary to at least a portion of a transcription initiation region of the newly synthesized RNA population;
 - digesting unhybridized, single-stranded RNA with a single-strand specific ribonuclease to form a protected RNA population;
 - detecting a protected RNA population comprising the labeled nucleotide analog; and
 - quantitating the protected RNA population comprising the labeled nucleotide analog.
17. The method of Claim 16, wherein contacting further comprises contacting with 2'-O-methyl-5-methyluridine-5'- triphosphate.

18. The method of Claim 16, further comprising providing the isolated replicase complex comprising the viral replicon template RNA by:

transfecting a human hepatosoma cell line with a viral replicon RNA or a DNA template for a viral replicon to provide a transfected cell line,

incubating the transfected cell line under conditions suitable for production of viral replicase complexes, and

isolating the replicase complexes comprising the viral replicon template RNA from the cell membrane fraction of the transfected cells.

19. The method of Claim 18, wherein the DNA template for the viral replicon is SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO:3, SEQ ID NO: 4, or SEQ ID NO: 5.

20. The method of Claim 16, further comprising providing the isolated replicase complex comprising the viral replicon template RNA by:

incubating a positive strand RNA virus infected primary cell or cell line under conditions suitable for production of viral replicase complexes, and

isolating the replicase complexes comprising the viral replicon template RNA from the cell membrane fraction of the infected cells.

21. The method of Claim 16, wherein the labeled nucleotide analog is an analog capable of being recognized by a specific antibody, an analog which can be recognized via a high specificity binding reaction, or an analog directly detectable as a result of a physical property of the analog.

22. The method of Claim 21, wherein the labeled nucleotide analog is an analog capable of being recognized by a specific antibody.

23. The method of Claim 22, wherein the labeled nucleotide analog is Br-UTP.

24. The method of Claim 22, wherein detecting comprises contacting the newly synthesized RNA with an antibody specific for the labeled nucleotide analog and immuno-precipitating the newly synthesized RNA comprising the labeled nucleotide analog to form immuno-precipitated RNA.

25. The method of Claim 24, wherein the labeled nucleotide analog is Br-UTP and the specific antibody is an anti-BrdU antibody.

26. The method of Claim 25, wherein quantitating the newly synthesized RNA comprises performing real time PCR on the immuno-precipitated RNA.

27. A method for determining whether a test compound is an RNA synthesis initiation inhibitor of a positive strand RNA virus comprising:

contacting an isolated replicase complex for the positive strand RNA virus comprising a viral replicon template RNA for the positive strand RNA virus, a labeled nucleotide analog, and the test compound, under conditions sufficient for *in vitro* RNA synthesis, to form a newly synthesized RNA population comprising the labeled nucleotide analog;

hybridizing a probe and the newly synthesized RNA population comprising the labeled nucleotide analog, under stringent hybridization conditions, wherein the probe is complementary to at least a portion of an initiation region of the newly synthesized RNA population;

digesting unhybridized, single-stranded RNA with a single-strand specific ribonuclease to form a protected RNA population;

detecting a protected RNA population comprising the labeled nucleotide analog;

quantitating the protected RNA population comprising the labeled nucleotide analog to provide a test RNA amount; and

comparing the test RNA amount with a control RNA amount of protected RNA comprising the labeled nucleotide analog but produced in the absence of the test compound, wherein a decrease in the test RNA amount compared to the control RNA amount indicates that the test compound inhibits RNA synthesis initiation of the positive strand RNA virus.

28. The method of Claim 27, wherein contacting further comprises contacting with a 2'-O-methyl-5-methyluridine-5'- triphosphate.

29. The method of Claim 27, further comprising providing the isolated replicase complex and the isolated viral replicon template RNA by:

transfecting a human hepatosoma cell line with a viral replicon RNA or a DNA template for a viral replicon to provide a transfected cell line,

incubating the transfected cell line under conditions suitable for production of viral replicase complexes, and

isolating the replicase complexes comprising the viral replicon template RNA from the cell membrane fraction of the transfected cells.

30. The method of Claim 29, wherein the DNA template for a viral replicon is SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO:3, SEQ ID NO: 4, or SEQ ID NO: 5.

31. The method of Claim 27, further comprising providing the isolated replicase complex comprising the viral replicon template RNA by:

incubating a positive strand RNA virus infected primary cell or cell line under conditions suitable for production of viral replicase complexes, and

isolating the replicase complexes comprising the viral replicon template RNA from the cell membrane fraction of the infected cells.

32. The method of Claim 27, wherein the positive strand RNA virus is Hepatitis C Virus.

33. The method of Claim 27, wherein the labeled nucleotide analog is an analog capable of being recognized by a specific antibody, an analog which can be recognized via a high specificity binding reaction, or an analog directly detectable as a result of a physical property of the analog.

34. The method of Claim 33, wherein the labeled nucleotide analog is an analog capable of being recognized by a specific antibody.

35. The method of Claim 34, wherein the labeled nucleotide analog is Br-UTP.

36. The method of Claim 34, wherein detecting comprises contacting the newly synthesized RNA with an antibody specific for the labeled nucleotide analog and immuno-precipitating the newly synthesized RNA comprising the labeled nucleotide analog to form immuno-precipitated RNA.

37. The method of Claim 36, wherein the labeled nucleotide analog is Br-UTP and the specific antibody is an anti-BrdU antibody.

38. The method of Claim 37, wherein quantitating the newly synthesized RNA comprises performing real time PCR on the immuno-precipitated RNA.

39. A kit for screening a test compound for inhibition of RNA synthesis of a positive strand RNA virus, comprising:

- a replicase complex for the positive strand RNA virus;
- a viral replicon template RNA for the positive strand RNA virus;
- instructions for use; and
- a buffer and nucleoside triphosphates sufficient for production of newly synthesized viral replicon RNA.

40. The kit of Claim 39, further comprising a labeled nucleotide analog.

41. The kit of Claim 40, further comprising a means for detecting the labeled nucleotide analog.

42. The kit of Claim 40, further comprising a means for quantitating the labeled nucleotide analog.

43. The kit of Claim 39, further comprising a primer, a single-strand specific ribonuclease, and instructions for performing a ribonuclease protection assay.

44. The kit of Claim 39, wherein the nucleoside triphosphates comprise 2'-O-methyl-5-methyluridine-5'- triphosphate.

QUANTITATIVE ASSAY FOR DETECTION OF NEWLY SYNTHESIZED RNA IN A
CELL-FREE SYSTEM AND IDENTIFICATION OF RNA SYNTHESIS INHIBITORS

ABSTRACT

[0089] Methods for detecting populations of newly synthesized and newly initiated RNAs in a cell-free system are described. Preferred templates for the populations of newly synthesized and newly initiated RNAs are viral replicon RNAs, particularly HCV viral replicon RNA. The newly synthesized and newly initiated RNAs may be formed in the presence of a labeled nucleotide analog suitable for detection and/or quantitation of the RNAs. The assay may be employed to identify small molecule inhibitors of RNA synthesis of positive strand RNA viruses such as Hepatitis C Virus.

RNase Protection Assay/Anti-BrdU purification/RT-PCR

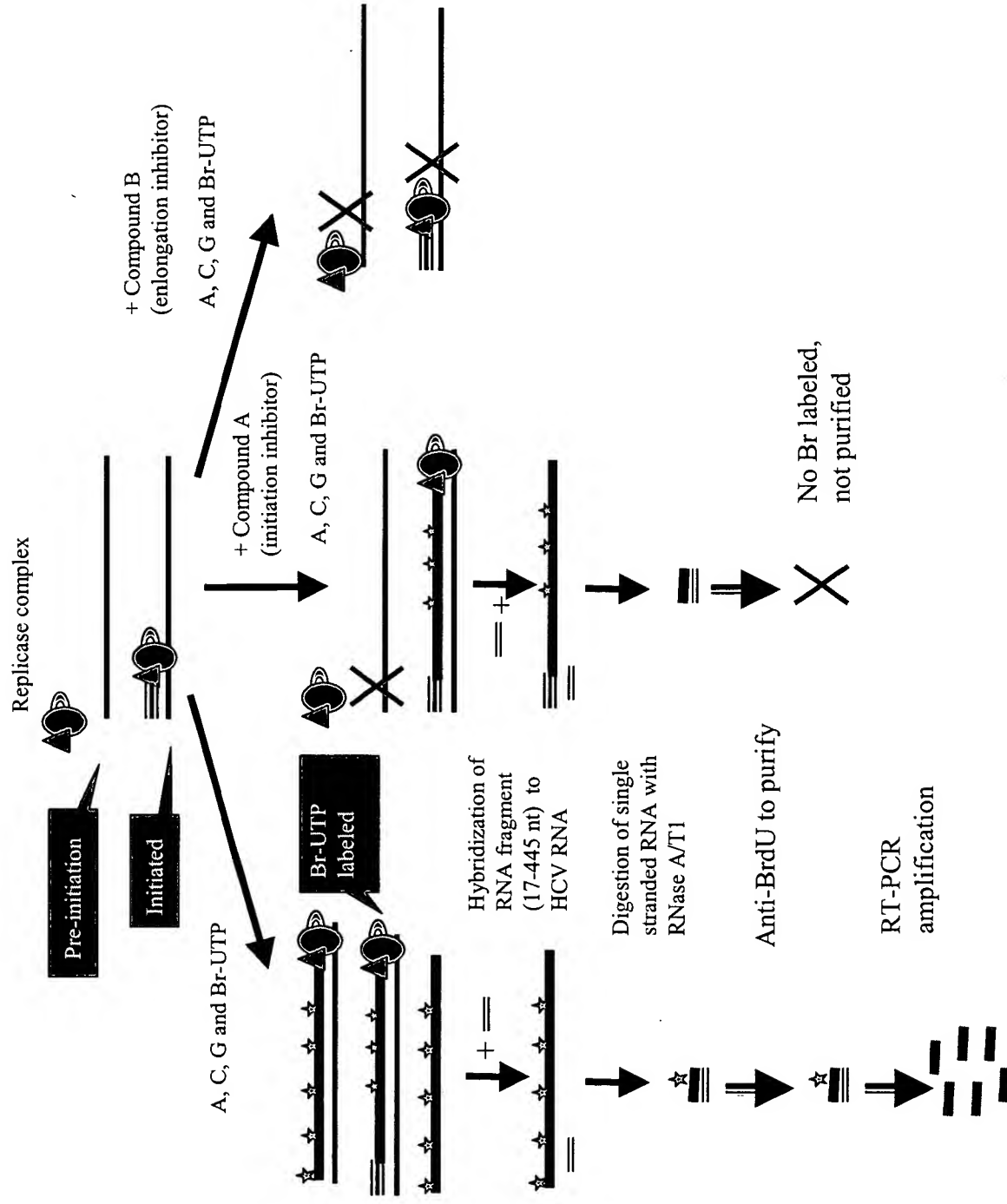
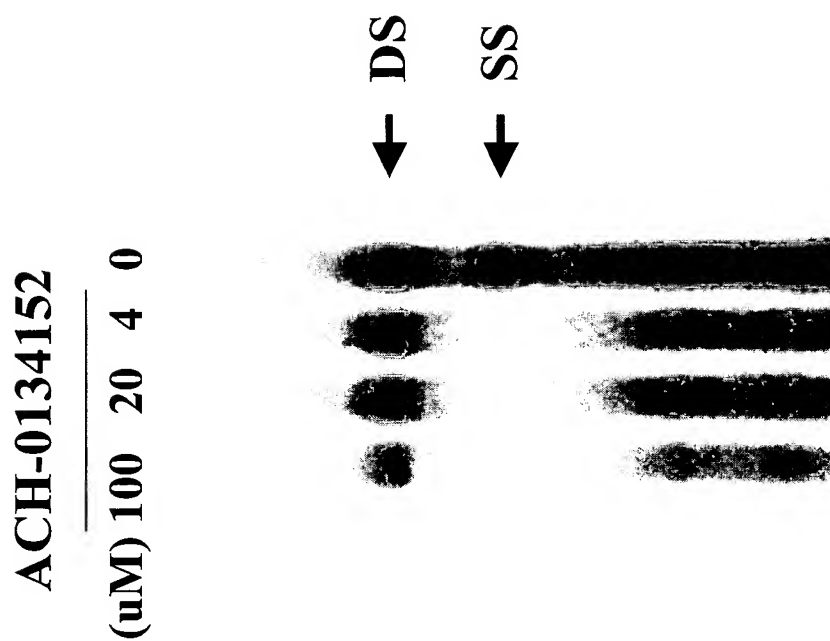
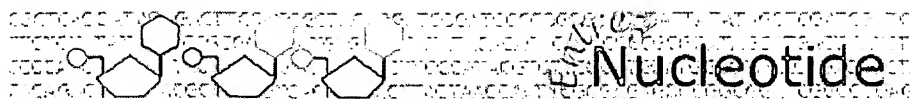
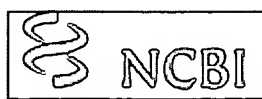


FIG. 1

FIG. 2



BEST AVAILABLE COPY



Entrez

PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

Br

Search for
1: [AJ242652](#). Hepatitis C virus...[gi:5441834]

Links

LOCUS SSE242652 7989 bp RNA linear SYN 11-MAY-2000
 DEFINITION Hepatitis C virus replicon I377/NS3-3'UTR.
 ACCESSION AJ242652
 VERSION AJ242652.1 GI:5441834
 KEYWORDS core-neo gene; NS2 gene; NS2 proteinase; NS3 gene; NS3
 proteinase/helicase; NS3/4A proteinase cofactor; NS4A gene; NS4B
 gene; NS4b protein; NS5a gene; NS5A phosphoprotein; NS5b gene; NS5B
 RNA dependant RNA polymerase; polyprotein.
 SOURCE Hepatitis C virus replicon I377/NS3-3'UTR
 ORGANISM [Hepatitis C virus replicon I377/NS3-3'UTR](#)
 artificial sequences; vectors.
 REFERENCE 1
 AUTHORS Lohmann,V., Korner,F., Koch,J., Herian,U., Theilmann,L. and
 Bartenschlager,R.
 TITLE Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell
 line
 JOURNAL Science 285 (5424), 110-113 (1999)
 MEDLINE [99322193](#)
 PUBMED [10390360](#)
 REFERENCE 2 (bases 1 to 7989)
 AUTHORS Bartenschlager,R.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAY-1999) Bartenschlager R., Institute for Virology,
 Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67,
 55131 Mainz, GERMANY
 FEATURES Location/Qualifiers
 source 1..7989
 /organism="Hepatitis C virus replicon I377/NS3-3'UTR"
 /mol_type="other RNA"
 /db_xref="taxon:95363"
 /focus
 /note="Neomycin-selectable bicistronic subgenomic
 hepatitis C virus replicon"
 source join(1..341,1801..7758)
 /organism="Hepatitis C virus"
 /mol_type="other RNA"
 /strain="1b"
 /isolate="Con1"
 /db_xref="taxon:11103"
 /tissue_type="liver"
 source 1190..1800
 /organism="Encephalomyocarditis virus"
 /mol_type="other RNA"
 /db_xref="taxon:12104"
 5'UTR 1..341
 /evidence=experimental
 gene 342..1181
 /gene="core-neo"
 CDS 342..1181
 /gene="core-neo"

```
/function="selectable marker"
/codon_start=1
/product="core-neo fusion protein"
/protein_id="CAB46912.1"
/db_xref="GI:5441835"
/translation="MSTNPKPQRKTKGRAMIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAVIDVVTGAGRDWLLLGEVPGQDLLSSHLAPAEKVSIMADAMRRLHTLDPATCFPDHQAHRIERARTRMEAGLVDQDDLDEEHQGLAPAEFLFARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRYQDIALATRDIAEELGGEWADRFLVLVYGIAAPDSQRIAFYRLLDEFF"
misc_feature 1190..1800
               /note="internal ribosome entry site (IRES)"
CDS           1801..7758
               /codon_start=1
               /product="non-structural polyprotein"
               /protein_id="CAB46913.1"
               /db_xref="GI:5441836"
               /translation="MAPITAYSQQTRGLLGCIITSLTGRDRNQVEGEVQVVSTATQSF
LATCVNGVCWTVYHGAGSKTLGAPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGS
SDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVC
TRGVAKAVDFVPVESMETTMRSPVFTDNSSPPAVPQTFQVAHLHAPTGSQKSTKVPAA
YAAQGYKVLVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTTGAPITYSTYQKFLA
DGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLAATATPPGSVTVP
PNIEEVALSSTGEIPFYGKAIPETIKGGRHLIFCHSKKKCDELAALKSLGLGLNAVAY
YRGLDVSVIPTSGDVIVVATDALMTGFTGDFDSDVCNTCVTQTVDFSLDPTFTIETT
TVPQDAVSRSQRRGRGTGRGRMGIYRFVTPGERPSGMFDSVLCYDAGCAWYELTPA
ETSVRLRAYLNTPLPVCQDHLFEWESVFTGLTHIDAHFLSQTQAGDNFPYLVAYQA
TVCARAQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGAVQNEVTTTHPITKYIMACM
SADLEVVTSTWVLVGGVLAALAAYCLTTGSSVIVGRIILSGKPAIIPDREVLVREFDE
MEECASHLPYIEQGMQLAEQFKQKAIGLLQTATKQAEAAAPVVESKWRTEAFWAKHM
WNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTQHTLLFNILGGWVAAQLAPP
SAASAFVGAGIAGAAVGSIGLGKVLVDILAGYGAGVAGALVAFKVMSEMPSTEDLVN
LLPAILSPGALVVGVCVCAAILRRHVGPGEAVQWMNRLIAFASRGNHVSPTHYVPESD
AAARVTQILSSLTITQLLKRHLHQWINECDSTPCSGSWLRDWDWICTVLTDFTKWLQS
KLLPRLPGVPPFFSCQRGYKGVWRGDGIMQTTCPGAQITGHVKNKSMRIVGPRTCSTN
WHGTFFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCP
CQVPAPEFFTEVDGVRHLHRYAPACKPLLREEVTFLVGLNQYLVSQLPCEPEPDVAVL
TSMLTDPSSHITAETAARLARGSPPSLASSASQLSAPSLKATCTTRHSDPADLIEA
NLLWRQEMGGNITRVESENKVILDSFEPLQAEEDEREVSVPAEILRRSRKFPRAMPI
WARPDYNPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPPRRKRTVVLSESTVSSALA
ELATKTFGSSSESAVDSGTATASPDQPSDDGDAGSDVESYSSMPLEGEPPDPLSDG
SWSTVSEEASEDVCCSMSYTWGTALITPCAAEETKLPIINALSNSLLRHHNLVYATTS
RSASLRQKKVTFDRLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSK
FGYGAKDVRNLSSKAVNHIRSVWKDLLEDTEPTIDTTIMAKNEVFVQPEKGRKPAR
LIVFPDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQRFELVNAWAKKCPMG
FAYDTRCFDSTVTENDIRVEESIQCCDLAPPEARQAIRSLTERLYIGGPLTNSKGQNC
GYRRCRASGVLTTSNGNTLTCYLKAAAACRAAKLQDCTMLVCGDDLVIICESAGTQED
EASLRAFTEAMTRYSAAPPDPPKPEYDLELITSCSSNSVVAHDASGRVYYLTRDPTT
PLARAAWETARHTPVNSWLGNIIYAPTLLWARMILMTHFFSILLAQEQLEKALDCQIY
GACYSIEPLDLPQIIQRLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARS
VRARLLSQGGRAATCGKYLFWAVRTKLKLTPIPAASQLDLSSWVAVAGYSGGDIYHSL
SRARPRWFMWCLLLLSVGVGIYLLPNR"
mat_peptide 1801..3696
               /gene="NS3"
               /product="NS3 proteinase/helicase"
mat_peptide 3697..3858
               /gene="NS4A"
               /product="NS3/4A proteinase cofactor"
mat_peptide 3859..4641
               /gene="NS4B"
               /product="NS4B protein"
mat_peptide 4642..5982
               /gene="NS5A"
```

```

        /product="NS5A phosphoprotein"
mat_peptide 5983..7755
              /gene="NS5B"
              /product="NS5B RNA dependant RNA polymerase"
gene         1801..3696
              /gene="NS3"
gene         3697..3858
              /gene="NS4A"
gene         3859..4641
              /gene="NS4B"
gene         4642..5982
              /gene="NS5A"
gene         5983..7755
              /gene="NS5B"
3'UTR       7759..7989
ORIGIN
1 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg
61 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggag
121 cccccctccc gggagagcca tagtggtctg cggaaccggg gagtacaccg gaattgccag
181 gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc
241 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg
301 gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac
361 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc
421 cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct
481 ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc gggtcttttt gtcaagaccg
541 acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca
601 cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc
661 tgctattggg cgaagtgcgc gggcaggatc tcctgtcatc tcaccttgct cctgccgaga
721 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc
781 cattcgacca ccaagcgaaa catcgcacgc agcgagcacg tactcggatg gaagccgggtc
841 ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg
901 ccaggctcaa ggcgcgcacg cccgacggcg aggatctcgt cgtgacctat ggcgatgcct
961 gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc
1021 tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc
1081 ttggcggcga atgggctgac cgcttctctg tgctttacgg tatcgccgct cccgattcgc
1141 agcgcacgca cttctatcgc cttcttgacg agttcttctg agtttaacaa gaccacaacg
1201 gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc
1261 cgaagccgct tggaaataagg ccggtgtgag tttgtctata tgttattttc caccatattg
1321 ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct
1381 aggggtcttt cccctctcgc caaaggaatg caaggctctg tgaatgtcgt gaaggaagca
1441 gttcctcttg aagcttcttg aagacaaaca acgtctgtag cgacctttg caggcagcgg
1501 aaccccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct
1561 gcaaaggcgg cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa
1621 tggctctcct caagcgtatt caacaagggg ctgaaggatg cccagaaggc acccattgt
1681 atgggatctg atctggggcc tcggtgcaca tgctttacat gtgtttagtc gaggttaaaa
1741 aacgtctagg cccccgaac cacggggacg tggttttcct ttgaaaaaca cgataatacc
1801 atggcgcta ttacggccta ctcccaacag acgcgaggcc tacttggtcg catcatcact
1861 agcctcacag gccgggacag gaaccagggtc gagggggagg tccaagtggc ctccaccgca
1921 acacaatctt tcctggcgac ctgcgtcaat ggcgtgtgtt ggactgtcta tcatggtgcc
1981 ggctcaaaga cccttgccgg cccaaagggc ccaatcacc aaatgtacac caatgtggac
2041 caggacctcg tcggctggca agcgcctccc ggggcgcgtt ccttgacacc atgcacctgc
2101 ggcagctcgg acctttactt ggtcacgagg catgccgatg tcattccggg gcgcggcgcg
2161 ggcgacagca gggggagcct actctcccc aggcgcgtct cctacttgaa gggctcttcg
2221 ggcggtccac tgctctgccc ctcggggcac gctgtgggca tctttcgggc tgccgtgtgc
2281 acccgagggg ttgcgaaggc ggtggacttt gtaccgctcg agtctatgga aaccactatg
2341 cggctcccgg tcttcacgga caactcgtcc cctccggccg taccgcagac attccagggtg
2401 gcccatctac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca
2461 gcccagggtg ataagggtgct tgtcctgaac ccgtccgtcg ccgccaccct aggtttcggg
2521 gcgtatatgt ctaaggcaca tggatcgac cctaacatca gaaccggggt aaggaccatc
2581 accacgggtg ccccatcac gtactccacc tatggcaagt ttcttgccga cgggtgttgc
2641 tctgggggcg cctatgacat cataatatgt gatgagtgcc actcaactga ctcgaccact
2701 atcctgggca tcggcacagt cctggaccaa gcggagacgg ctggagcgcg actcgtcgtg
2761 ctgcgccagg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggagggtg
2821 gctctgtcca gcaactggaga aatccccctt tatggcaaag ccatccccat cgagaccatc

```

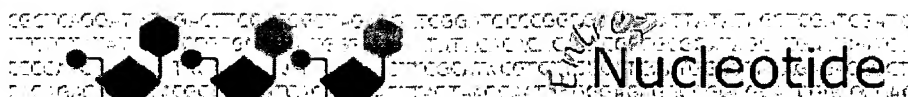
```
2881 aagggggggga ggcacctcat tttctgccat tccaagaaga aatgtgatga gctcgccgcg
2941 aagctgtccg gcctcggact caatgctgta gcatattacc ggggccttga tgtatccgctc
3001 ataccaacta gcggagacgt cattgtcgta gcaacggacg ctctaatagac gggctttacc
3061 ggcgatttcg actcagtgat cgactgcaat acatgtgtca cccagacagt cgacttcagc
3121 ctggacccga ccttcacat tgagacgacg accgtgccac aagacgcggt gtcacgctcg
3181 cagcggcgag gcaggactgg taggggcagg atgggcattt acaggtttgt gactccagga
3241 gaacggccct cgggcatggt cgattcctcg gttctgtgcg agtgctatga cgcgggctgt
3301 gcttggtacg agctcacgcc cgccgagacc tcagttagggt tgcgggctta cctaaacaca
3361 ccagggttgc ccgtctgcca ggaccatctg gaggttctggg agagcgtctt tacaggcctc
3421 accacacatg acgcccattt cttgtcccag actaagcagg caggagacaa ctcccctac
3481 ctggtagcat accaggctac ggtgtgcgag agggctcagg ctccacctcc atcgtgggac
3541 caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gcccttgctg
3601 tataggctgg gagccgttca aaacgagggt actaccacac accccataac caaatacatc
3661 atggcatgca tgtcggctga cctggagggt gtcacgagca cctgggtgct ggtaggcgga
3721 gtctagcag ctctggccgc gtattgcctg acaacaggca gcgtggtcat tgtgggcagg
3781 atcatcttgt ccggaaagcc ggccatcatt cccgacaggg aagtccttta ccgggagttc
3841 gatgagatgg aagagtgcgc ctcacacctc cttacatcg aacagggaat gcagctcgcc
3901 gaacaattca aacagaaggc aatcgggttg ctgcaaacag ccaccaagca agcggagggt
3961 gctgtcccg tggtggaatc caagtggcgg accctcgaag ccttctgggc gaagcatatg
4021 tggaaatttca tcagcgggat acaatattta gcaggcttgt ccactctgcc tggcaacccc
4081 gcgatagcat cactgatggc attcacagcc tctatcacca gcccgctcac caccacat
4141 accctcctgt ttaacatcct ggggggatgg gtggccgccc aacttgctcc tccagcgct
4201 gcttctgctt tcgtaggcgc cggcatcgct ggagcggctg ttggcagcat aggccttggg
4261 aagggtgctt tggatatttt ggcagggtat ggagcagggg tggcaggcgc gctcgtggcc
4321 ttaaggtca tgagcggcga gatgccctcc accgaggacc tggttaacct actccctgct
4381 atcctctccc ctggcgccct agtcgtcggg gtctgtgctg cagcgatact gcgtcggcac
4441 gtgggcccag gggagggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgagg
4501 ggtaaccacg tctccccac gcactatgtg cctgagagcg acgctgcacg actgtcact
4561 cagatcctct ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcaac
4621 gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatatgc
4681 acggtgttga ctgatttcaa gacctggctc cagtcacaag tcttgcgcgc attgcccggg
4741 gtccccctct tctcatgtca acgtgggtac aagggagttc ggcggggcga cggcatcatg
4801 caaaccacct gcccatgtgg agcacagatc accggacatg tgaaaaacgg ttccatgagg
4861 atcgtggggc ctaggacctg tagtaacacg tggcatggaa cattccccat taacgcgtac
4921 accacggggc cctgcacgcc ctccccggcg ccaaattatt ctagggcgtg gtggcgggtg
4981 gctgctgagg agtacgtgga ggttacgcgg gtgggggatt tccactacgt gacgggcatg
5041 accactgaca acgtaaagtg cccgtgtcag gttccggccc ccgaattctt cacagaagtg
5101 gatgggggtg ggttgcacag gtacgtcca gcgtgcaaac cctcctacg ggaggaggtc
5161 acattcctgg tcgggtctaa tcaataacct gttgggtcac agtcccatg cgagcccga
5221 ccggacgtag cagtgtcac tccatgtct accgacctc cccacattac ggcggagacg
5281 gctaagcgta ggctggccag gggatctccc cctccttgg ccagctcatc agctagccag
5341 ctgtctgcgc ctctcttgaa ggcaacatgc actaccgctc atgactcccc ggacgctgac
5401 tcatcgagg ccaacctcct gtggcggcag gagatgggcg ggaacatcac ccgctggag
5461 tcagaaaata aggtagtaat tttgactct tccagccgc tccaagcgca ggaggatgag
5521 agggagatat ccgttcgggc ggagactctg cggaggcca ggaaattccc tcgagcgatg
5581 cccatatggg cagccccgga ttacaacct ccaactgttag agtcttgaa ggaccggac
5641 tacgtccctc cagtgttaca cgggtgtcca ttgccgctg ccaaggcccc tccgatacca
5701 cctccacgga ggaagaggac ggttgtcctg tcagaatcta ccgtgtcttc tgccttggcg
5761 gagctcgcca caaagacctt cggcagctcc gaatcgctcg ccgtcgacag cggcacggca
5821 acggcctctc ctgaccagcc ctccgacgac ggcgacgcgg gatccgacgt tgagtcgtac
5881 tctccatgc ccccccttga gggggagccg ggggatcccg atctcagcga cgggtcttgg
5941 tctaccgtaa gcgaggaggc tagtgaggac gtctgtgct gctcgatgtc ctacacatg
6001 acaggccccc tgatcacgcc atgcgtgcg gaggaaacca agctgccat caatgcactg
6061 agcaactctt tgcctcgta ccacaacttg gtctatgcta caacatctcg cagcgcaagc
6121 ctgcggcaga agaagggtcac ctttgacaga ctgcaggctc tggacgacca ctaccgggac
6181 gtgtcaagg agatgaaggc gaaggcgtcc acagttaagg ctaaaactct atccgtggag
6241 gaagcctgta agctgacgcc cccacattcg gccagatcta aatttggcta tggggcaag
6301 gagtcggga acctatccag caaggccgtt aaccacatcc gctccgtgtg gaaggacttg
6361 ctggaagaca ctgagacacc aattgacacc accatcatgg caaaaaatga ggttttctgc
6421 gtccaaccag agaagggggg ccgcaagcca gctcgctta tegtattccc agatttggg
6481 gttcgtgtgt gcgagaaaat ggccctttac gatgtggtct ccacctccc tcaggccgtg
6541 atgggctctt catacggatt ccaatactct cctggacagc gggctcaggtt cctgggtaat
6601 gcctggaaag cgaagaaatg ccctatgggc ttgcgatatg acaccgcgtg ttttgactca
6661 acggtcactg agaatgacat ccgtgttgag gagtcaatct accaatgttg tgacttggcc
```

```
6721 cccgaagcca gacaggccat aaggtcgctc acagagcggc tttacatcgg gggccccctg
6781 actaattcta aagggcagaa ctgcggctat cgccggtgcc gcgcgagcgg tgtactgacg
6841 accagctgcg gtaataccct cacatgttac ttgaaggccg ctgcggcctg tcgagctgcg
6901 aagctccagg actgcacgat gctcgtatgc ggagacgacc ttgtcgttat ctgtgaaagc
6961 gcggggaccc aagaggacga ggcgagccta cgggccttca cggaggctat gactagatac
7021 tctgcccccc ctggggaccc gcccaaacca gaatacgact tggagttgat aacatcatgc
7081 tcctccaatg tgtcagtcgc gcacgatgca tctggcaaaa ggggtgtacta tctcaccctg
7141 gacccccacca ccccccttgc gcgggctgcg tgggagacag ctagacacac tccagtcaat
7201 tcctggctag gcaacatcat catgtatgcg ccaccttgt gggcaaggat gatcctgatg
7261 actcatttct tctccatcct tctagctcag gaacaacttg aaaaagccct agattgtcag
7321 atctacgggg cctgttactc cattgagcca cttgacctac ctcagatcat tcaacgactc
7381 catggcctta gcgcattttc actccatagt tactctccag gtgagatcaa taggggtggct
7441 tcatgcctca ggaaacttgg ggtaccgcc ttgcgagtct ggagacatcg ggccagaagt
7501 gtccgcgcta ggctactgtc ccaggggggg agggctgcc cttgtggcaa gtacctcttc
7561 aactgggcag taaggaccaa gctcaaactc actccaatcc cggctgcgtc ccagttggat
7621 ttatccagct ggttcgttgc tggttacagc gggggagaca tatatcacag cctgtctcgt
7681 gcccgaaccc gctggttcat gtggtgccta ctctacttt ctgtaggggt aggcatttat
7741 ctactcccca accgatgaac ggggagctaa aactccagg ccaataggcc atcctgtttt
7801 tttccctttt tttttttctt tttttttttt tttttttttt tttttttttt ttctcctttt
7861 tttttcctct ttttttcctt ttctttcctt tggtggtcc atcttagccc tagtcacggc
7921 tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc
7981 agatcaagt
```

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Mar 15 2004 07:22:16



Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy Bi

Search for

Limits Preview/Index History Clipboard Details
 Show:

☐ 1: [AB114136](#). Hepatitis C virus...[gi:40714444] [Links](#)

LOCUS AB114136 8024 bp DNA linear SYN 27-JAN-2004
 DEFINITION Hepatitis C virus replicon pSGR-JFH1 gene for neomycin resistance gene product, hepatitis C virus nonstructural protein, complete cds.
 ACCESSION AB114136
 VERSION AB114136.1 GI:40714444
 KEYWORDS .
 SOURCE Hepatitis C virus replicon pSGR-JFH1
 ORGANISM Hepatitis C virus replicon pSGR-JFH1
 artificial sequences; vectors.
 REFERENCE 1
 AUTHORS Kato,T., Date,T., Miyamoto,M., Furusaka,A., Tokushige,K., Mizokami,M. and Wakita,T.
 TITLE Efficient Replication of Hepatitis C Virus Genotype 2a Subgenomic Replicon
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 8024)
 AUTHORS Kato,T. and Wakita,T.
 TITLE Direct Submission
 JOURNAL Submitted (04-JUL-2003) Takanobu Kato, Nagoya City University Graduate School of Medical Sciences, Department of Clinical Molecular Informative Medicine; 1, Kawasumi, Mizuho-ku,, Nagoya, Aichi 467-8601, Japan (E-mail:ta_kato@med.nagoya-cu.ac.jp, Tel:81-52-853-8292, Fax:81-52-842-0021)
 FEATURES Location/Qualifiers
 source 1..8024
 /organism="Hepatitis C virus replicon pSGR-JFH1"
 /mol_type="other DNA"
 /db_xref="taxon:237617"
 /focus
 source 1..397
 /organism="Hepatitis C virus"
 /mol_type="other DNA"
 /isolation_source="IH/fulminant hepatitis"
 /db_xref="taxon:11103"
 /clone="JFH-1"
 /country="Japan:Tokyo"
 /note="5'UTR
 genotype:2a"
 source 1207..1767
 /organism="Encephalomyocarditis virus"
 /mol_type="other DNA"
 /db_xref="taxon:12104"
 source 1774..8024
 /organism="Hepatitis C virus"
 /mol_type="other DNA"
 /isolation_source="IH/fulminant hepatitis"
 /db_xref="taxon:11103"
 /clone="JFH-1"
 /country="Japan:Tokyo"

gene /note="NS3 - 3'X
genotype:2a"
398..1192
/gene="NEO"
CDS 398..1192
/gene="NEO"
/codon_start=1
/product="neomysin resistance gene product"
/protein_id="BAD06941.1"
/db_xref="GI:40714445"
/translation="MIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGR
PVL FVKTDLSGALNELQDEAARLSWLATTGPVCAAVLDVVTEAGRDWLLLGEVPGQDL
LSSHLAPAEKVSIMADAMRRLHTLDPATCFPDHQA KHRIERARTRMEAGLVDQDDLDE
EHQGLAPAE L FARLKARMPDGEDLVVTHGDA CLPNIMVENGRFSGFIDCGRLGVADRY
QDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLDEFF"
misc_feature 1207..1767
/note="IRES sequence"
gene 1774..7788
/gene="NS3-NS5b"
CDS 1774..7788
/gene="NS3-NS5b"
/codon_start=1
/product="hepatitis C virus nonstructural protein"
/protein_id="BAD06942.1"
/db_xref="GI:40714446"
/translation="MAPITAYAQQTRGLLGAIVVSMTGRDRTEQAGEVQILSTVSQS F
LGT TISGVLWTVYHGAGNKTLAGLRGPVTQMYSSAEGDLVGWSPSPGTKSLEPCKCGA
VDLYLVTRNADVIPARRRGDKRGALLSPRP ISTLKGSSGGPVLCPRGHVVGLFRAAVC
SRGVAKSIDFIPVETLDVVTRSP TFS DNSTPPAVPQTYQVGYLHAPTGS GKSTKVPVA
YAAQGYKVLVLNPSVAATLGFGAYLSKAHGINPNIRTGVRTVMTGEAITYSTY GKFLA
DGGC ASGAYDIIICDECHAVDATSILGIGTVLDQAETAGVRLTVLATATPPGSVTTTPH
PDIEEVGLGREGEIPFYGRAIPLSCIKGGRHLIFCHSKKKCDELAALRGMGLNAVAY
YRGLDVSIIIPAQGDVVVVATDALMTGYTGDFDSVIDCNVAVTQAVDFS LDPTFTITTO
TVPQDAVSRSQRRGRTGRGRQGT YRYVSTGERASGMFDSVVLCECYDAGAAWYDLTPA
ETTVRLRAYFNTPLGPVCQDHLEFWEAVFTGLTHIDAHFLSQT KQAGENFAYLVAYQA
TVCARAKAPPPSWDAMWKCLARLKPTLAGPTPLLYRLGPITNEVTLTHPGTKYIATCM
QADLEVMTSTWVLAGGVLA AVAAYCLATGCVSIIIGRLHVNQRVVVAPDKEVLYEAFDE
MEECASRAALIEEGORIAEMLKSKIQGLLQQASKQAQDIQPAMQASWPKEQFWARHM
WNFISGIQYLAGLSTLPGNPAVASMMAFSAALTSPLSTSTTILLNIMGGWLASQIAPP
AGATGFVVVSGLVGA AVGSIGLGKVLVDILAGYGAGISGALVAFKIMSGEKPSMEDVIN
LLPGILSPGALVVGVICAAILRRHVGPGE GAVQWMNRLIAFASRGNHVAPTHYVTESD
ASQRTVQLLGLSTITSLRLRLHNWITEDCPIPCSGSWLRD VWDVWCTILTD FKNWLTS
KLFPKLPGLPFIS CQKGYKGVWAGTGIMTTRCPCGANISGNVRLSGMRTITGPKTCMNT
WQGTFPINCYTEGQCAPKPPTNYKTAIWRVAASEYAEVTQHGSYSYVTGLTTDNLKIP
CQLPSPEFFSWVDGVQIHRFAPTPKPF RDEVSCFVGLNSYAVGSQLPCEPEPDADV L
RSM L TDP PHITAE TAARRLARGSP PSEASSSVS QLSAPSLRATCTTHSNTYDVMVDA
NLLMEGGVAQTEPESRVPVLD FLEPMAEEESDLEPSIPSECMLPRSGFPRALPAWARP
DYNPPLVESWRRPDYQPPTVAGCALPPPKAPT PPRRRRTVGLSESTISEALQQLAI
KTFGQPPSSGDAGSSTGAGAAESGGPTSPGEPAPSETGSASSMPPLEGEPGDPDLESD
QVELQPPPQGGGVAPGSGSGSWSTCSEEDTTVCCSMSYSWTGALITPCSPEEEKLP I
NPLSN SLLRYHNKVYCTTSKSASQRAKKVTFDR TQVLD AHYDSVLKDIKLAASKVSAR
LLTLEEACQLTPPHSARSKYGF GAKEVRSLSGRAVNHIKSVWKD LLED PQTPIPTTIM
AKNEVFCVDPAKGKKPARLIVY PDLGVRVCEKMALYDITQKLPQAVMGASYGFQYSP
AQRVEYLLKAWAEKKDPMGFSYDTRCFDSTVTERDIRTEESIYQACSLPEEARTAIHS
LTERLYVGGPMFNSKGQTCGYRRCRASGVLT TSMGNTITCYVKALAACKAAGIVAPTM
LVCGDDL VVISESQGTEEDERNLRAFTEAMTRY SAPPGDPPRPEYDLELITSCSSNVS
VALGPRGRRRYL TRDPTT PLARA AWETVRHSPINSWLGNI IQYAPT IWVRMVLMT HF
FSILMVQDTLDQNLNFEMYGSVYSVNPLDLPAIIERLHGLDAFSMHTYSHHELTRVAS
ALRKL GAPPLRVWKSRAVRASLISRGGKA AVCGRYLFNWA VKTKLKLTLPEARLL
DLSSWFTV GAGGGDIFHSVSRARPRSL LFGLLLLFVGVLFLLPAR"

ORIGIN

1 acctgccct aatagggcg acactccgcc atgaatcact cccctgtgag gaactactgt
61 cttcacgcag aaagcgcta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc
121 ccccctcccg ggagagccat agtggctctgc ggaaccggtg agtacaccgg aattgccggg

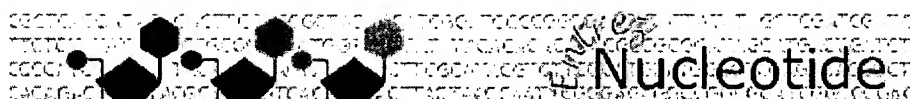
```
181 aagactgggt cctttcttgg ataaacccac tctatgcccg gccatttggg cgtgccccg
241 caagactgct agccgagtag cgttgggttg cgaaaggcct tgttgtagtg cctgataggg
301 cgcttgcgag tgccccggga ggtctcgtag accgtgcacc atgagcacia atcctaacc
361 tcaaagaaaa accaaaagaa acaccaaccg tcgccaatg attgaacaag atggattgca
421 cgcaggttct ccggccgctt ggggtggagag gctattcggc tatgactggg cacaacagac
481 aatcggctgc tctgatgccg ccgtgttccg gctgtcagcg caggggcgcc cggttctttt
541 tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag gacgaggcag cgcggctatc
601 gtggctggcc acgacgggcg ttccttgccg agctgtgctc gacgttgtca ctgaagcggg
661 aagggactgg ctgctatttg gcgaagtgcc ggggcaggat ctctgtcatc ctacacttgc
721 tccctgcgag aaagtatcca tcatggctga tgcaatgcgg cggctgcata cgcttgatcc
781 ggctacctgc ccattcgacc accaagcgaa acatcgcatc gagcgagcac gactcggat
841 ggaagccggg cttgtcgatc aggatgatct ggacgaagag catcaggggc tcgcgccagc
901 cgaactgttc gccaggtcca aggcgcgcac gcccgacggc gaggatctcg tcgtgacca
961 tggcgatgcc tgcttgccga atatcatggt ggaaaatggc cgcttttctg gattcatcga
1021 ctgtggcccg ctgggtgtgg cggaccgcta tcaggacata gcgttggcta cccgtgatat
1081 tgctgaagag cttggcggcg aatgggctga ccgcttcctc gtgctttacg gtatcgccgc
1141 tcccgattcg cagcgcatcg ccttctatcg ccttcttgac gaggttctct gaggtttaac
1201 cctctccctc cccccccctt aacgttactg gccgaagccg cttggaataa ggccggtgtg
1261 cgtttgtcta tatgttattt tccaccatat tgccgtcttt tggcaatgtg agggcccgga
1321 aacctggccc tgtctctctg acgagcattc ctagggggtc tccccctctc gccaaaggaa
1381 tgcaaggtct gttgaatgtc gtgaagggaag cagttctctc ggaagcttct tgaagacaaa
1441 caacgtctgt agcgaccctt tgcaggcagc ggaaccccc accctggcgac aggtgcctct
1501 gcggccaaaa gccacgtgta taagatacac ctgcaaaggc ggcacaacc cagtgccacg
1561 ttgtgagttg gatagttgtg gaaagagtca aatggctctc ctcaagcgta ttcaacaagg
1621 ggctgaagga tgcccagaag gtaccccat gtatgggata tgatctgggg cctcgggtga
1681 catgctttat atgtgtttag tcgaggttaa aaaaacgtct agggcccccg aaccacgggg
1741 acgtggtttt cctttgaaaa acacgatgat accatggctc ccactactgc ttatgcccag
1801 caaacacgag gcctcctggg cgccatagtg gtgagtatga cggggcgtag caggacagaa
1861 caggccgggg aagtccaaat cctgtccaca gtctctcagt ccttctcgg aacaaccatc
1921 tcgggggttt tgtggactgt ttaccacgga gctggcaaca agactctagc cggcttacgg
1981 ggtccgggtc cgcagatgta ctcgagtgtc gagggggact tggtaggtcg gccagcccc
2041 cctgggacca agtcttttga gccgtgcaag tgtggagccg tcgacctata tctggtcacg
2101 cggaacgctg atgtcatccc ggctcggaga cgcggggaca agcggggagc attgctctcc
2161 ccgagaccca ttctgacctt gaaggggtcc tcgggggggc cgggtgctct ccctaggggc
2221 cactcgcttg ggctcttccg agcagctgtg tgctctcggg gcgtggccaa atccatcgat
2281 ttcaccccgc ttgagacact cgcgtttgtt acaaggctc ccactttcag tgaacagac
2341 acgccaccgg ctgtgcccc aacctatcag gtcgggtact tgcatgctcc aactggcagt
2401 ggaaagagca ccaaggctcc tgtcgcgtat gccgccagg ggtacaaagt actagtgtt
2461 aaccctcgg tagctgccac cctgggggtt ggggcgtacc tatccaaggc acatggcatc
2521 aatcccaaca ttaggactgg agtcaggacc gtgatgaccg gggaggccat cactactcc
2581 acatatggca aatttctcgc cgatgggggc tgcgctagcg gcgcctatga catcatcata
2641 tgcgatgaat gccacgctgt ggatgtacc tccattctcg gcatcggaac ggtccttgat
2701 caagcagaga cagccgggg cagactaact gtgctggcta cggccacacc cccgggtca
2761 gtgacacccc cccatcccga tatagaagag gtaggcctcg ggcgggaggg tgagatccc
2821 ttctatggga gggcgattcc cctatctcgc atcaaggag ggagacacct gattttctgc
2881 cactcaaaga aaaagtgtga cgagctcgcg gcggcccttc ggggcatggg cttgaatgcc
2941 gtggcatact atagaggggt ggacgtctcc ataataccag ctcagggaga tgtggtggtc
3001 gtcgccaccg acgcccctat gacgggggtac actggagact ttgactccgt gatcgactgc
3061 aatgtagcgg tcaccaagc tgtcgacttc agcctggacc ccaccttcac tataaccaca
3121 cagactgtcc cacaagacgc tgtctcacgc agtcagcgc gcgggcgcac aggtagagga
3181 agacagggca cttataggtat gttttccact ggtgaacgag cctcaggaa gtttgacagt
3241 gtagtgtttt gtgagtgtat cgcagcagg gctgcgtggg acgatctcac accagcagag
3301 accacgctca ggcttagagc gtatttcaac acgcccggcc taccggtgtg tcaagaccat
3361 cttgaatttt gggaggcagt tttcacccgc ctacacaca tagacgcca cttcctctcc
3421 caaacaagc aagcggggga gaacttcgcg tacctagtag cctaccaagc tacggtgtgc
3481 gccagagcca agggccctcc ccgctcctgg gacgccatgt ggaagtgcct ggccgactc
3541 aagcctacgc ttgcggggcc cacacctctc ctgtaccgtt tgggccttat taccaatgag
3601 gtcaccctca cacaccctgg gacgaagtac atcgccacat gcatgcaagc tgaccttgag
3661 gtcacgacca gcacgtgggt cctagctgga ggagtcctgg cagccgtcgc cgcatttgc
3721 ctggcgactg gatgcgtttc catcatcggc cgcttgacag tcaaccagcg agtcgtcgtt
3781 gcgcgggata aggaggtcct gtatgaggt tttgatgaga tggaggaaat gcctctagg
3841 gcggtcttca tcgaagaggg gcagcgggata gccgagatgt tgaagtccaa gatccaaggc
3901 ttgctgcagc aggcctctaa gcaggcccag gacatacaac ccgctatgca ggcttcatgg
3961 cccaaagtgg aacaattttt ggccagacac atgtggaact tcattagcgg catccaatac
```

```
4021 ctgcgaggat tgtcaacact gccagggaac cccgcggtgg cttccatgat ggcattcagt
4081 gccgccctca ccagtcogtt gtcgaccagt accaccatcc ttctcaacat catgggaggc
4141 tgggttagcgt ccagatcgc accaccgcg ggggccaccg gctttgtcgt cagtggcctg
4201 gtgggggctg ccgtgggcag cataggcctg ggtaagggtg tggtagacat cctggcagga
4261 tatggtgcgg gcatttcggg ggccctcgtc gcattcaaga tcatgtctgg cgagaagccc
4321 tctatggaag atgtcatcaa tctactgcct gggatcctgt ctccgggagc cctgggtggtg
4381 ggggtcatct gcgcggccat tctgcgcgcg cacgtgggac cgggggaggg cgcggtccaa
4441 tggatgaaca ggcttattgc ctttgcttcc agaggaaacc acgtcgcccc tactactac
4501 gtgacggagt cggatgcgtc gcagcgtgtg acccaactac ttggctctct tactataacc
4561 agcctactca gaagactcca caattggata actgaggact gccccatccc atgtccgga
4621 tcctggctcc gcgacgtgtg ggactgggtt tgcaccatct tgacagactt caaaaattgg
4681 ctgacctcta aattgttccc caagctgccc ggccctccct tcatctcttg tcaaaagggg
4741 tacaagggtg tgtgggcggg cactggcctc atgaccacgc gctgcccttg cggcgccaac
4801 atctctggca atgtccgcct gggctctatg aggatcacag ggcctaaaac ctgcatgaac
4861 acctggcagg ggacctttcc tatcaattgc tacacggagg gccagtgcgc gccgaaaccc
4921 cccacgaact acaagaccgc catctggagg gtggcggcct cggagtacgc ggaggtgacg
4981 cagcatgggt cgtactccta tgtaacagga ctgaccactg acaatctgaa aattccttgc
5041 caactacctt ctccagagtt tttctcctgg gtggacgggt tgcagatcca taggtttgca
5101 cccacaccaa agccgttttt ccgggtagag gtctcgttct gcgttggggt taattcctat
5161 gctgtcgggg cccagcttcc ctgtgaacct gagcccgacg cagacgtatt gaggtccatg
5221 ctaacagatc cgccccacat cacggcggag actgcggcgc ggcgcttggc acggggatca
5281 cctccatctg aggcgagctc ctcatgagc cagctatcag caccgtcgct gcggggccacc
5341 tgcaccaccc acagcaacac ctatgacgtg gacatggtcg atgccaacct gctcatggag
5401 ggcggtgtgg ctacagacaga gcctgagtcg aggggtgccg ttctggactt tctcgagcca
5461 atggccgagg aagagagcga ccttgagccc tcaataacct cggagtgcac gctccccagg
5521 agcgggtttc cacgggcctt accggttgg gcacggcctg actacaaccc gcgctcgtg
5581 gaatcgtgga ggaggccaga ttaccaacgg cccaccgttg ctggttgtgc tctccccccc
5641 cccaagaagg ccccgacgcc tcccccaagg agacgcccga cagtgggtct gatcgagagc
5701 accatatcag aagccctcca gcaactggcc atcaagacct ttggccagcc cccctcgagc
5761 ggtgatgcag gctcgtccac gggggcgggc gccgcgaat ccggcggtcc gacgtccctt
5821 ggtgagccgg cccctcaga gacaggttcc gcctcctcta tgccccccct cgaggggggag
5881 cctggagatc cggacctgga gtctgatcag gtagagcttc aacctcccc ccagggggggg
5941 ggggtagctc ccggttcggg ctccgggtct tgggtactt gctccgagga ggacgatacc
6001 accgtgtgct gctccatgtc atactcctgg accggggctc taataactcc ctgtagcccc
6061 gaagaggaaa agttgccaat caacccttgg agtaactcgc tgttgcgata ccataacaag
6121 gtgtactgta caacatcaaa gagcgcctca cagagggtca aaaaggtaac ttttgacagg
6181 acgcaagtgc tcgacgcccc ttatgactca gtcttaaagg acatcaagct agcggcttcc
6241 aaggtcagcg caaggctcct caccttgagg gaggcgtgcc agttgactcc accccattct
6301 gcaagatcca agtatggatt cggggccaag gaggtccgca gcttgtccgg gagggccggt
6361 aaccacatca agtcggtgtg gaaggacctc ctggaagacc cacaacacc aattcccaca
6421 accatcatgg ccaaaaatga ggtgtctctg gtggaccccc ccaagggggg taagaaacca
6481 gctcgcctca tcgtttaccc tgacctcggc gtccgggtct gcgagaaaat ggccctctat
6541 gacattacac aaaagcttcc tcaggcggtg atgggagctt cctatggctt ccagtactcc
6601 cctgcccaac gggtaggata tctcttgaaa gcatgggcgg aaaagaagga cccctgggt
6661 ttttcgtatg ataccgatg ctctgactca accgtcactg agagagacat caggaccgag
6721 gagtccatat accaggcctg ctccctgccc gaggaggccc gcactgccat acactcgctg
6781 actgagagac ttacgtagg agggcccatg ttcaacagca agggtcaaac ctgcggttac
6841 agacgttgcc gcgccagcgg ggtgctaacc actagcatgg gtaacaccat cacatgctat
6901 gtgaaagccc tagcggcctg caaggctgcg gggatagttg cgcgccaat gctggtatgc
6961 ggcgatgacc tagtagtcat ctacagaaagc caggggactg agggaggacga gcggaacctg
7021 agagccttca cggaggccat gaccaggtag tctgcccctc ctggtgatcc cccagaccg
7081 gaatatgacc tggagctaat aacatcctgt tctcaaatg tgtctgtggc gttgggcccg
7141 cggggccgcc gcagatacta cctgaccaga gacccaacca ctccactcgc cggggctgcc
7201 tgggaaacag ttagacactc cctatcaaat tcatggctgg gaaacatcat ccagtatgct
7261 ccaaccatat gggttcgcac ggtcctaata acacacttct tctccattct catggtccaa
7321 gacaccctgg accagaacct caactttgag atgtatggat cagtatactc cgtgaatcct
7381 ttggaccttc cagccataat tgagagggtta cacgggcttg acgccttttc tatgcacaca
7441 tactctcacc acgaactgac gcgggtgggt tcagccctca gaaaacttgg ggcgccaccc
7501 ctcagggtgt ggaagagtcg ggctcgcgca gtcagggcgt cctcatctc ccgtggaggg
7561 aaagcggccg tttgcggccg atatctcttc aattgggcgg tgaagaccaa gctcaaactc
7621 actccattgc cggaggcgcg cctactggag ttatccagtt ggttcacgt cgggcgggc
7681 gggggcgaca tttttcacag cgtgtcgcgc gcccgacccc gctcattact ctccggccta
7741 ctccactttt tcgtaggggg aggcctcttc ctactccccg ctcggtagag cggcacacac
7801 taggtacact ccatagctaa ctgttccttt tttttttttt tttttttttt
```

```
7861 tttttttttt cttttttttt tttttccctc ttttttccct tctcatctta ttctactttc
7921 tttcttggtg gctccatctt agccctagtc acggctagct gtgaaaggtc cgtgagccgc
7981 atgactgcag agagtgccgt aactggtctc tctgcagatc atgt
//
```

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Mar 15 2004 07:22:16



Entrez PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

Bi

Search for

1: [AJ242654](#). Hepatitis C virus...[gi:5441840]

Links

LOCUS SSE242654 8001 bp RNA linear SYN 11-MAY-2000
 DEFINITION Hepatitis C virus replicon I389/NS3-3'UTR.
 ACCESSION AJ242654
 VERSION AJ242654.1 GI:5441840
 KEYWORDS core-neo gene; NS2 gene; NS2 proteinase; NS3 gene; NS3 proteinase/helicase; NS3/4A proteinase cofactor; NS4A gene; NS4B gene; NS4b protein; NS5a gene; NS5A phosphoprotein; NS5b gene; NS5B RNA dependant RNA polymerase; polyprotein.
 SOURCE Hepatitis C virus replicon I389/NS3-3'UTR
 ORGANISM Hepatitis C virus replicon I389/NS3-3'UTR
 artificial sequences; vectors.
 REFERENCE 1
 AUTHORS Lohmann,V., Korner,F., Koch,J., Herian,U., Theilmann,L. and Bartenschlager,R.
 TITLE Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line
 JOURNAL Science 285 (5424), 110-113 (1999)
 MEDLINE 99322193
 PUBMED 10390360
 REFERENCE 2 (bases 1 to 8001)
 AUTHORS Bartenschlager,R.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAY-1999) Bartenschlager R., Institute for Virology, Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67, 55131 Mainz, GERMANY
 FEATURES . Location/Qualifiers
 source 1..8001
 /organism="Hepatitis C virus replicon I389/NS3-3'UTR"
 /mol_type="other RNA"
 /db_xref="taxon:95365"
 /focus
 /note="Neomycin-selectable bicistronic subgenomic hepatitis C virus replicon"
 source join(1..341,1813..7770)
 /organism="Hepatitis C virus"
 /mol_type="other RNA"
 /strain="1b"
 /isolate="Con1"
 /db_xref="taxon:11103"
 /tissue_type="liver"
 /country="Germany"
 source 1202..1812
 /organism="Encephalomyocarditis virus"
 /mol_type="other RNA"
 /db_xref="taxon:12104"
 5'UTR 1..341
 /evidence=experimental
 gene 342..1193
 /gene="core-neo"
 CDS 342..1193

/gene="core-neo"
/function="selectable marker"
/codon_start=1
/product="core-neo fusion protein"
/protein_id="CAB46916.1"
/db_xref="GI:5441841"
/translation="MSTNPKPQRKTKRNTNGRAMIEQDGLHAGSPAAWVERLFGYDWA
QQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAALVDV
VTEAGRDWLLLGEPVPGQDLLSSHLAPAEKVSIMADAMRRLHTLDPATCFPDHQAKHRI
ERARTRMEAGLVDQDDLDEEHQGLAPAELEFARLKARMPDGEDLVVTHGDACLPNIMVE
NGRFSGFIDCGRLGVADRYQDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFY
RLLDEFF"
misc_feature 1202..1812
/note="internal ribosome entry site (IRES)"
CDS 1813..7770
/codon_start=1
/product="non-structural polyprotein"
/protein_id="CAB46917.1"
/db_xref="GI:5441842"
/translation="MAPITAYSQQTRGLLGCIITSLTGRDRNQVEGEVQVSTATQSF
LATCVNGVCWTVYHGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGS
SDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVC
TRGVAKAVDFVPVESMETTMRSPVFTDNSSPPAVPQTFQVAHLHAPTGSGBKSTKVPAA
YAAQGYKVLVLNPSVAATLGFAYMSKAHGIDPNIRTGVRTITTTGAPITYSTYKFLA
DGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGARLVVLAATATPPGSVTVP
PNIEEVALSSTGEIPFYGAIPETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAY
YRGLDVSVIPTSGDIVVATDALMTGTGDFDSVIDCNTCVTQTVDFSLDPTFTIETT
TVPQDAVSRSQRRGRTGRGRMGIIYRFVTPGERPSGMFDSVLCCEYDAGCAWYELTPA
ETSVRLRAYLNTPLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQA
TVCARAQAPPPSWDQMWKCLIRLKP TLHGPTLLYRLGAVQNEVTTTHPI TKYIMACM
SADLEVVTSTWVLVGGVLAALAAAYCLTTGSSVIVGRIILSGKPAIIPDREVLVREFDE
MEECASHLPYIEQGMQLAEQFKQKAIGLLQTATKQAEAAAPVVESKWRTEAFWAKHM
WNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTQHTLLFNILGGWVAAQLAPP
SAASAFVGAGIAGAAVGSIGLGKVLVDILAGYGAGVAGALVAFKVMSEMPSTEDLVN
LLPAILSPGALVVGVCVCAAILRRHVGPGEAVQWMNRLIAFASRGNHVSPTHYVPESD
AAARVTQILSSLTITQLLKRLHQWINEDECSTPCSGSWLRDWDWICTVLTDFTKTLWQS
KLLPRLPGVPPFFSCQRGYKGVWRGDGIMQTTCPGCAQITGHVKNGSMRIVGPRTCSNT
WHGTFPINAYTTGPCTPSPAPNYSRALWRVAAEEYVEVTRVGD FHYVTGMTTDNVKCP
CQVPAPEFFTEVDGVRHLHRYAPACKPLLREEVTFLVGLNQYL VGSQLPCEPEPDVAVL
TSMLTDP SHITAETA KRRLARGSPPSLASSASQLSAPSLKATCTTRHDS PADLIEA
NLLWRQEMGGNITRVESENKVVILDSFEPLQAEEDEREVSVP AEILRRSRKFPRAMP
IWARPDYNPPLLESWKDPDYVPPVHGCPLPPAKAPP IPPRRKRTVVLSSESTVSSALA
ELATKTFGSSSESAVDSGTATASPDQPSDDGDAGSDVESYSSMPPLEGEPGDPDLSDG
SWSTVSEEA SEDVCCSMSYTWGALITPCAAEETKLPINALSNSLLRHHNLVYATTS
RSASLRQKKVTFDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLT PPHSARSK
FGYGAKDVRLNLSKAVNHIRSVWKDLLED TETPIDTTIMAKNEVFCVQPEKGRKPAR
LIVFPDLGVRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQRFELVNAWKAKKCPMG
FAYDTRCFDSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNC
GYRRCRASGVLTTSCGNTLTCYLKAAAACRAAKLQDCTMLVCGDDL VVICESAGTQED
EASLRAFTEAMTRYSAAPGDPPKPEYDLELITSCSSNVSVAH DASGRVYYLTRDPTT
PLARAAWETARHTPVNSWLGNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIY
GACYSIEPLDLPQIIQRLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARS
VRARLLSQGGRAATCGKYLEFNWAVRTKLKLTPIPAASQLDLSSWVFVAGYSGGDIYHSL
SRARPRWFMWCLLLLSVGVGIYLLPNR"
mat_peptide 1813..3708
/gene="NS3"
/product="NS3 proteinase/helicase"
mat_peptide 3709..3870
/gene="NS4A"
/product="NS3/4A proteinase cofactor"
mat_peptide 3871..4653
/gene="NS4B"
/product="NS4B protein"
mat_peptide 4654..5994

```

/ gene="NS5A"
/ product="NS5A phosphoprotein"
mat_peptide 5995..7767
/ gene="NS5B"
/ product="NS5B RNA dependant RNA polymerase"
gene 1813..3708
/ gene="NS3"
gene 3709..3870
/ gene="NS4A"
gene 3871..4653
/ gene="NS4B"
gene 4654..5994
/ gene="NS5A"
gene 5995..7767
/ gene="NS5B"
3'UTR 7771..8001

```

ORIGIN

```

1 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg
61 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac
121 cccccctccc gggagagcca tagtgggtct cggaaccggg gagtacaccg gaattgccag
181 gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc
241 gcgagactgc tagccgagta gtgttggtgc gcgaaaggcc ttgtggtact gcctgatagg
301 gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc
361 ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc
421 acgcagggtc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga
481 caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt
541 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat
601 cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg
661 gaagggactg gctgctattg ggcgaaagtgc cggggcagga tctcctgtca tctcaccttg
721 ctctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc
781 cggctacctg ccatttcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga
841 tgggaagccg tcttgctgat caggatgatc tggacgaaga gcatcagggg ctgcgcgcag
901 ccgaactgtt cgccaggctc aaggcgcgca tgcccgcagg cgaggatctc gtcgtgacct
961 atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg
1021 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata
1081 ttgctgaaga gcttgccggc gaatgggctg accgcttctc cgtgctttac ggtatcgccg
1141 ctcccgattc gcagcgcata gccttctatc gccttcttga cgagttcttc tgagtttaaa
1201 cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct
1261 aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt
1321 tccaccatat tgccgtcttt tggcaatgtg agggcccggg aacctggccc tgtctcttg
1381 acgagcattc ctaggggtct tccccctctc gccaaaggaa tgcaagggtc gttgaatgtc
1441 gtgaagggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt
1501 tgcaggcagc ggaaccccc accctggcagc aggtgcctct gcggccaaaa gccacgtgta
1561 taagatacac ctgcaaggc ggacacaccc cagtgccacg ttgtgagttg gatagttgtg
1621 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag
1681 gtacccatt gtatgggatc tgatctgggg cctcgggtgca catgctttac atgtgtttag
1741 tcgagggttaa aaaacgtcta ggccccccga accacgggga cgtggttttc ctttgaaaaa
1801 cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc
1861 tgcatacatc ctagcctcac aggcggggac aggaaccagg tcgaggggga ggtccaagtg
1921 gtctccaccg caacacaatc tttcctggcg acctgcgtca atggcggtg ttggactgtc
1981 tatcatggtg ccggctcaaa gaccttgcc ggcccaaagg gcccaatcac ccaaattgtac
2041 accaatgttg accaggacct cgtcggtcgg caagcgcccc ccggggcgcg ttccttgaca
2101 ccatgcacct gcggcagctc ggacctttac ttggtcacga ggcatgccga gttcttccg
2161 gtgcgcgggc ggggcgacag caggggggagc ctactctccc ccaggcccgt ctctacttg
2221 aagggtctct cgggcggtcc actgctctgc ccctcggggc acgctgtggg catctttcgg
2281 gctgcggtgt gcacccgagg ggttgccaag gcggtggact ttgtaccctg cgagtctatg
2341 gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag
2401 acattccagg tggcccatct acacgcccc actggtagcg gcaagagcac taagggtgcg
2461 gctgcgtatg cagcccaagg gtataagggt cttgtcctga acccgctcgt cgccgccacc
2521 ctagggtttc gggcgatatg gtctaaggca catggtatcg accctaactc cagaaccggg
2581 gtaaggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gtttcttgcc
2641 gacggtggtt gctctggggg cgccatgac atcataatat gtgatgagtg ccaactcaact
2701 gactcgacca ctatcctggg catcggcaca gtctggacc aagcggagac ggctggagcg
2761 cgactcgctc tgctcgccac cgctacgctc ccgggatcgg tcaccgtgcc acatccaaac

```



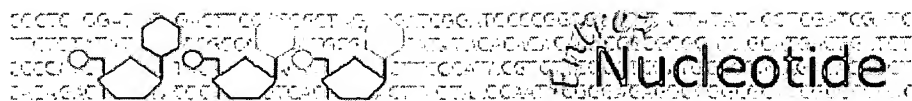
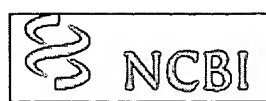
```
2821 atcgaggagg tggctctgtc cagcactgga gaaatccctt tttatggcaa agccatcccc
2881 atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat
2941 gagctcgccg cgaagctgtc cggcctcgga ctcaatgctg tagcatatta cgggggcctt
3001 gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaattg
3061 acgggcttta ccggcgattt cgactcagtg atcgactgca atacatgtgt caccagaca
3121 gtcgacttca gcctggaccc gaccttcacc attgagacga cgaccgtgcc acaagacgcg
3181 gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt
3241 gtgactccag gagaacggcc ctccggcatg ttcgattcct cggttctgtg cgagtgttat
3301 gacgcgggct gtgcttggtg cgagctcacg cccgccgaga cctcagttag gttgcgggct
3361 tacctaaaca caccagggtt gccgctctgc caggaccatc tggagttctg ggagagcgtc
3421 ttacaggcc tcaccacat agacgcccac ttcttgccc agactaagca ggcaggagac
3481 aacttcccc acctggtagc ataccaggct acggtgtgcg ccagggtcga ggctccacct
3541 ccatcgtagg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacgggcca
3601 acgccccctg tgtataggct gggagccggt caaacgagg ttactaccac acaccccata
3661 accaaatata tcatggcatg catgtcggct gacctggagg tcgtcacgag cactgggtg
3721 ctggtaggcg gagtcctagc agctctggcc gcgtattgcc tgacaacagg cagcgtggtc
3781 attgtgggca ggatcatctt gtccggaaa ggcgccatca tccccgacag ggaagtcctt
3841 taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga
3901 atgcagctcg ccgaacaatt caaacagaag gcaatcgggt tgctgcaaac agccaccaag
3961 caagcggagg ctgctgctcc cgtgggtgaa tccaagtggc ggacctcga agccttctgg
4021 gcgaagcata tgtggaattt catcagcggg atacaatatt tagcaggctt gtccactctg
4081 cctggcaacc ccgcgatagc atcactgatg gcattcacag cctctatcac cagcccgctc
4141 accacccaac ataccctcct gttaacatc ctggggggat ggggtggcgc ccaacttgct
4201 cctcccagcg ctgcttctgc tttcgtaggc gccggcatcg ctggagcggc tgttggcagc
4261 ataggccttg ggaagggtgt tgtggatatt ttggcagggt atggagcagg ggtggcaggc
4321 gcgctcgtgg cctttaaggc catgagcggc gagatgccct ccaccgagga cctggttaac
4381 ctactccctg ctatcctctc cctggcgcc ctagtctgct gggctcgtgt cgcagcgata
4441 ctgcgtcggc acgtggggcc aggggaggg gctgtgcagt ggatgaaccg gctgatagcg
4501 ttcgcttcgc ggggtaacca cgtctcccc acgcactatg tgctgagag cgacgctgca
4561 gcacgtgtca ctcagatcct ctctagtctt accatcactc agctgctgaa gaggcttcac
4621 cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgtttgg
4681 gattggatat gcacgggtgt gactgatttc aagacctggc tccagtccaa gctcctgccg
4741 cgattgcccg gagtccccct ctctcatgt caacgtgggt acaagggagt ctggcggggc
4801 gacggcatca tgcaaaccac ctgcccatgt ggagcacaga tcaccggaca tgtgaaaaac
4861 ggttccatga ggatcgtggg gcctaggacc tgtagtaaca cgtggcatgg aacattcccc
4921 attaacgctg acaccacggg ccctgcacg ccctccccg cgccaaatta ttctaggcg
4981 ctgtggcggg tggctgctga ggagtagctg gaggttacgc ggggtggggg tttccactac
5041 gtgacgggca tgaccactga caacgtaaa tgcccgtgtc aggttccggc ccccgaaattc
5101 ttcacagaag tggatggggg gcggttgac aggtacgctc cagcgtgcaa acccctccta
5161 cgggaggagg tcacattcct ggtcgggctc aatcaatacc tggttgggtc acagctccca
5221 tgcgagcccg aaccggacgt agcagtgtct acttccatgc tcaccgaccc ctcccacatt
5281 acggcggaga cggctaagcg taggctggcc aggggatctc cccctcctt ggccagctca
5341 tcagctagcc agctgtctgc gccttccttg aaggcaacat gcactacccg tcatgactcc
5401 ccggacgctg acctcatcga ggccaacctc ctgtggcggc aggagatggg cgggaacatc
5461 acccgctgg agtcagaaaa taaggtagta attttgact ctttcgagcc gctccaagcg
5521 gaggaggatg agagggaaat atccgttccg gcggagatcc tgccggaggtc caggaaattc
5581 cctcgagcga tgcccatatg ggcacgccc gattacaacc ctccactgtt agagtctgg
5641 aaggaccccg actacgtccc tccagtggta cacgggtgtc cattgccgcc tgccaaggcc
5701 cctccgatac cacctccacg gaggaagagg acggttgtcc tgtcagaatc taccgtgtct
5761 tctgccttgg cggagctcgc cacaagacc ttcggcagct ccgaatcgtc ggccgtcgac
5821 agcggcacgg caacggcctc tcctgaccag ccctccgacg acggcgacgc gggatccgac
5881 gttgagtcgt actcctccat gccccccctt gagggggagc cgggggatcc cgatctcagc
5941 gcagggtcct ggtctaccgt aagcgaggag gctagtgagg acgtcgtctg ctgctgatg
6001 tcctacacat ggacaggcgc cctgatcacg ccatgcgctg cggaggaaac caagctgccc
6061 atcaatgcac tgagcaactc tttgtctcgt caccacaact tggcttatgc tacaacatct
6121 cgcagcgcaa gcctgcggca gaagaaggct accttgaca gactgcaggc cctggacgac
6181 cactaccggg acgtgctcaa ggagatgaag gcgaaggcgt ccacagttaa ggctaaactt
6241 ctatccgtgg aggaagcctg taagctgacg cccccacatt cggccagatc taaatttggc
6301 tatggggcaa aggacgtccg gaacctatcc agcaaggccg ttaaccacat ccgctccgtg
6361 tggaaaggact tgctggaaga cactgagaca ccaattgaca ccaccatcat ggcaaaaaat
6421 gaggttttct gcgtccaacc agagaagggg ggcgcgaagc cagctcgctt tatcgatttc
6481 ccagatattg gggttcgtgt gtgcgagaaa atggccctt acgatgtggg ctccacctc
6541 cctcaggccg tgatgggctc ttcatacgga ttccaatact ctctgggaca gcgggtcgag
6601 ttcttggtga atgcctggaa agcgaagaaa tgccctatgg gcttcgcata tgacacccgc
```

```
6661 tgttttgact caacgggtcac tgagaatgac atccgtgttg aggagtcaat ctaccaatgt
6721 tgtgacttgg cccccgaagc cagacaggcc ataaggctgc tcacagagcg gctttacatc
6781 ggggggcccc tgactaattc taaagggcag aactgcggct atcgccggtg ccgcgcgagc
6841 ggtgtactga cgaccagctg cggtaatacc ctcacatgtt acttgaaggc cgctgcggcc
6901 tgtcgagctg cgaagctcca ggactgcacg atgctcgtat gcggagacga ccttgtcgtt
6961 atctgtgaaa gcgcggggac ccaagaggac gaggcgagcc tacgggcctt cacggaggct
7021 atgactagat actctgcccc ccctggggac ccgcccacac cagaatacga cttggagtgt
7081 ataacatcat gtcctcccaa tgtgtcagtc gcgcacgatg catctggcaa aaggggtgtac
7141 tatctcaccg gtgacccacac cccccctt gcgcgggctg cgtgggagac agctagacac
7201 actccagtca attcctggct aggcacatc atcatgtatg cggccacctt gtgggcaagg
7261 atgacctga tgactcattt cttctccatc cttctagctc aggaacaact tgaaaaagcc
7321 ctagattgtc agatctacgg ggctgttac tccattgagc cacttgacct acctcagatc
7381 attcaacgac tccatggcct tagcgcattt tactccata gttactctcc aggtgagatc
7441 aatagggtgg cttcatgcct caggaaactt ggggtaccgc ctttgcgagt ctggagacat
7501 cgggccagaa gtgtccgcgc taggtactg tcccaggggg ggagggtgc cacttgtggc
7561 aagtacctct tcaactgggc agtaaggacc aagctcaaac tcaactccat cccggctgcg
7621 tcccagttgg atttatccag ctggttcgtt gctggttaca gcgggggaga catatatcac
7681 agcctgtctc gtgcccagac ccgctggttc atgtggtgcc tactcctact ttctgtaggg
7741 gtaggcacat atctactccc caaccgatga acggggagct aaacactcca ggccaatagg
7801 ccatcctgtt tttttccctt ttttttttc ttttttttt ttttttttt ttttttttt
7861 ttttctcctt ttttttccct ctttttttcc ttttctttcc tttggtggct ccatcttagc
7921 cctagtcacg gctagctgtg aaaggctccg gagccgcttg actgcagaga gtgctgatac
7981 tggcctctct gcagatcaag t
```

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Mar 15 2004 07:22:16



Entrez

PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

Bi

Search

Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

Display

default

Show:

20

Send to

File

Get Subsequence

Feat

1: [AJ242653](#). Hepatitis C virus...[gi:5441837]

Links

LOCUS SSE242653 8649 bp RNA linear SYN 11-MAY-2000
 DEFINITION Hepatitis C virus replicon I389/NS2-3'UTR.
 ACCESSION AJ242653
 VERSION AJ242653.1 GI:5441837
 KEYWORDS core-neo gene; NS2 gene; NS2 proteinase; NS3 gene; NS3
 proteinase/helicase; NS3/4A proteinase cofactor; NS4A gene; NS4B
 gene; NS4b protein; NS5a gene; NS5A phosphoprotein; NS5b gene; NS5B
 RNA dependant RNA polymerase; polyprotein.
 SOURCE Hepatitis C virus replicon I389/NS2-3'UTR
 ORGANISM Hepatitis C virus replicon I389/NS2-3'UTR
 artificial sequences; vectors.
 REFERENCE 1
 AUTHORS Lohmann,V., Korner,F., Koch,J., Herian,U., Theilmann,L. and
 Bartenschlager,R.
 TITLE Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell
 line
 JOURNAL Science 285 (5424), 110-113 (1999)
 MEDLINE 99322193
 PUBMED 10390360
 REFERENCE 2 (bases 1 to 8649)
 AUTHORS Bartenschlager,R.
 TITLE Direct Submission
 JOURNAL Submitted (26-MAY-1999) Bartenschlager R., Institute for Virology,
 Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67,
 55131 Mainz, GERMANY
 FEATURES Location/Qualifiers
 source 1..8649
 /organism="Hepatitis C virus replicon I389/NS2-3'UTR"
 /mol_type="other RNA"
 /db_xref="taxon:95364"
 /focus
 /note="Neomycin-selectable bicistronic subgenomic
 hepatitis C virus replicon"
 source join(1..341,1813..8418)
 /organism="Hepatitis C virus"
 /mol_type="other RNA"
 /strain="1b"
 /isolate="Con1"
 /db_xref="taxon:11103"
 /tissue_type="liver"
 /country="Germany"
 source 1202..1812
 /organism="Encephalomyocarditis virus"
 /mol_type="other RNA"
 /db_xref="taxon:12104"
 5'UTR 1..341
 /evidence=experimental
 gene 342..1193
 /gene="core-neo"
 CDS 342..1193

```
/gene="core-neo"
/function="selectable marker"
/codon_start=1
/product="core-neo fusion protein"
/protein_id="CAB46914.1"
/db_xref="GI:5441838"
/translation="MSTNPKPQRKTKRNTNGRAMIEQDGLHAGSPAAWVERLFGYDWA
QQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPAAVLDV
VTEAGRDWLLLGEVPGQDLLSSHLAPAEKVSIMADAMRRLHTLDPATCFDQAKHRI
ERARTRMEAGLVDQDDLDDEHQGLAPAEFLARLKARMPDGEDLVVTHGDACLPNIMVE
NGRFSGFIDCGRLGVADRYQDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFY
RLLDEFF"
misc_feature 1202..1812
               /note="internal ribosome entry site (IRES)"
CDS           1813..8418
               /codon_start=1
               /product="non-structural polyprotein"
               /protein_id="CAB46915.1"
               /db_xref="GI:5441839"
               /translation="MDREMAASCGGAVFVGLILLTLSPHYKLFLARLIWWLQYFITRA
EAHLQVWIPPLNVRGGRDAVILLTCAIHPELIFTITKILLAILGPLMVLQAGITKVPY
FVRAHGLIRACMLVRKVAGGHYVQMALMKLAALTGTYYVDHLTPLRDWAHAGLRDLAV
AVEPVVFSDMETKVIWGTADTAACGDIILGLPVSARRGREIHLGPADSLEGQGWRLLA
PITAYSQQTRGLLGCIITSLTGRDRNQVEGEVQVSTATQSFLATCVNGVCWTVYHGA
GSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYLVTRHADVIPVR
RRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVESM
ETTMRSVPVFTDNSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVA
ATLGFGAYMSKAHGIDPNIRTGVRTITTGAPITYSTYKFLADGGCSGGAYDIIICDE
CHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSSTGEIPF
YGKAIPETIKGGRHLIFCHSKKKCDELA AKLSGLGLNAVAYYRGLDVSVIPTSGDVI
VVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGRT
GRGRMGIIYRFVTPGERPSGMFSSVLC ECYDAGCAWYELTPAETS VRLRAYLNT PGLP
VCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAPPPSWDQM
WKCLIRLKP TLHGPTPLLYRLGAVQNEVTTHPITKYIMACMSADLEVVTSTWVLVGG
VLAALAAAYCLTTGSVVIVGRIILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQ
LAEQFKQKAIGLLQTATKQAEAAAPVVESKWR TLEAFWAKHMWNFISGIQYLAGLSTL
PGNPAIASLMAFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASAFVGAGIAGAAV
GSIGLGKVLVDILAGYGAGVAGALVAFKVMSEMPSTEDLVNLLPAILSPGALVVG VV
CAAILRRHVGPGE GAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTQILSSLTITQ
LLKRLHQWINE DCSTPCSGSWLRD VWDWICTVLTD FKTWLQSKLLPRLPGVPFFSCQR
GYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIVGPRTCSNTWHGTFFPINAYTTGPCT
PSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPQVPAP EFFFTEVDGVR
LHRYAPACKPLLREEVTFVLVGLNQYLVGSQLPCEPEPDVAVLTSMLTDP SHITAETAK
RRLARGSPPSLASSSASQLSAPSLKATCTTRHDS PDADLIEANLLWRQEMGGNITRVE
SENKVVIDLSFEPLQAEEDEREVSVP AEIILRRSRKFPRAMPIWARPDYNPPLLESWKD
PDYVPPVVHGCPLPPAKAPP IPPRRKRRTVVLSESTVSSALAE LATKTFGSSESAVD
SGTATASPDQPSDDGDAGSDVESYSSMPPLEGEPGDPDLS DGSWSTVSE EASEDVVCC
SMSYTWGTALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASLRQKKVTFDRLQ
VLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKF GYGAKDVRLNLSKAV
NHIRSVWKDLLED TETPIDTTIMAKNEVF CVQPEKGGRKPARLIVFPDLGVRVCEKMA
LYDVVSTLPQAVMGSSYGFQYSPGQ RVEFLVNAWKAKKCPMGFAYDTRCFDSTVTEND
IRVEESYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTS CG
NTLTCYLKAAAACRAAKLQDCTMLVCGDDL VVICESAGTQEDEASLRAFTEAMTRYSA
PPGDPPKPEYDLELITSCSSNVSVAH DASGKR VYYLTRDPTT PLARA AWETARHTPVN
SWLGNIIMYAPTLWARMILMTHFFS ILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQ
RLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRATCG
KYLFWNAVRTKLKLTPIPAASQLDLSSWFVAGYSGGDIYHSLSRARPRWFMWCLLLLL
VGVGIYLLPNR"
mat_peptide 1813..2463
               /gene="NS2"
               /product="NS2 proteinase"
mat_peptide 2464..4356
               /gene="NS3"
               /product="NS3 proteinase/helicase"
```

mat_peptide 4357..4518
 /gene="NS4A"
 /product="NS3/4A proteinase cofactor"
mat_peptide 4519..5301
 /gene="NS4B"
 /product="NS4B protein"
mat_peptide 5302..6642
 /gene="NS5A"
 /product="NS5A phosphoprotein"
mat_peptide 6643..8415
 /gene="NS5B"
 /product="NS5B RNA dependant RNA polymerase"
gene 1813..2463
 /gene="NS2"
gene 2464..4356
 /gene="NS3"
gene 4357..4518
 /gene="NS4A"
gene 4519..5301
 /gene="NS4B"
gene 5302..6642
 /gene="NS5A"
gene 6643..8415
 /gene="NS5B"
3'UTR 8419..8649

ORIGIN

```

1 gccagccccc gattggggggc gacactccac catagatcac tcccctgtga ggaactactg
61 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac
121 cccccctccc gggagagcca tagtgggtctg cggaaccggt gagtacaccg gaattgccag
181 gacgaccggg tcttttcttg gatcaaccgg ctcaatgcct ggagatttgg gcgtgcccc
241 gcgagactgc tagccgagta gtgttggtgc gcgaaaggcc ttgtggtact gcctgatagg
301 gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc
361 ctcaaagaaa aaccaaaccg aacaccaacg ggcgcgccat gattgaacaa gatggattgc
421 acgcagggtc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga
481 caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggggcg ccggttcttt
541 ttgtcaagac cgacctgtcc ggtgcctga atgaactgca ggacgaggca gcgcggctat
601 cgtggctggc cagcaggggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg
661 gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg
721 ctctgcccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc
781 cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga
841 tgggaagcgg tcttgctgat caggatgatc tggacgaaga gcatcagggg ctgcgcgacg
901 ccgaactgtt cgccaggctc aaggcgcgca tgcccgaagg cgaggatctc gtcgtgaccc
961 atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg
1021 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggtc acccgtgata
1081 ttgctgaaga gcttggcggc gaattgggctg accgcttcct cgtgctttac ggtatcgccg
1141 ctcccgattc gcagcgcata gccttctatc gccttcttga cgagttcttc tgagtttaaa
1201 cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct
1261 aacgttactg gccgaagcgg cttggaataa ggccgggtgt cgtttgtcta tatgttattt
1321 tccaccatat tgccgtcttt tggcaatgtg agggcccggg aacctggccc tgtcttcttg
1381 acgagcattc ctaggggtct tccccctctc gccaaaggaa tgcaaggctc gttgaatgtc
1441 gtgaagggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt
1501 tgcaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta
1561 taagatacac ctgcaaaggc gccacaaccc cagtgccacg ttgtgagttg gatagttgtg
1621 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag
1681 gtacccatt gtatgggatc tgatctgggg cctcgggtgca catgctttac atgtgtttag
1741 tgcagggttaa aaaacgtcta ggccccccga accacgggga cgtggtttct ctttgaaaaa
1801 cacgataata ccatggaccg ggagatggca gcatcgtgcg gaggcgcggg tttcgtaggt
1861 ctgatactct tgaccttgtc accgcactat aagctgttcc tcgctaggct catatggtgg
1921 ttacaatatt ttatcaccag ggccgagga cacttgcaag tgtggatccc cccctcaac
1981 gttcgggggg gccgcgatgc cgtcacctc ctcacgtgcg cgatccacc agagctaatc
2041 ttaccatcca ccaaactctt gctcgccata ctcggtccac tcatgggtgc ccaggctggg
2101 ataaccaaa tgccgtactt cgtgcgcgca caggggtca ttcgtgcatg catgctgggtg
2161 cggaaggttg ctgggggtca ttatgtccaa atggctctca tgaagttggc cgcactgaca
2221 ggtacgtacg tttatgacca tctcacccca ctgcgggact gggcccacgc gggcctacga
  
```

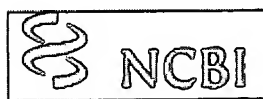
```
2281 gaccttgccg tggcagttga gcccgtcgtc ttctctgata tggagaccaa ggttatcacc
2341 tgggggggcag acaccggcgc gtgtggggac atcatcttgg gectgcccgt ctccgcccgc
2401 aggggggaggg agatacatct gggaccggca gacagccttg aagggcaggg gtggcgactc
2461 ctgcgccta ttacggccta ctcccaacag acgcgagggc tacttggtg catcatcact
2521 agcctcacag gccgggacag gaaccaggtc gagggggagg tccaagtggg ctccaccgca
2581 acacaatctt tcctggcgac ctgcgtcaat ggcgtgtgtt ggactgtcta tcatgggtgc
2641 ggctcaaaga cccttgccgg cccaaagggc ccaatcaccc aatgtacac caatgtggac
2701 caggacctcg tcggctggca agcgcctccc ggggcgcgtt ccttgacacc atgcacctgc
2761 ggcagctcgg acctttactt ggtcacgagg catgccgatg tcattccggg gcgcggcggg
2821 ggcagacaga gggggagcct actctcccc actgccgtct cctacttgaa gggctcttcg
2881 ggcgggtocac tgctctgccc ctccggggcac gctgtgggca tctttcgggc tgcgtgtgc
2941 acccgagggg ttgcgaaggc ggtggacttt gtaccgctcg agtctatgga aaccactatg
3001 cggctccccg tcttcacgga caactcgtcc cctccggccg taccgcagac attccagggtg
3061 gcccatctac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca
3121 gcccaagggt ataagggtgt tgcctgaac ccgtccgtcg ccgccacct aggtttcggg
3181 gcgtatatgt ctaaggcaca tggtagcgc cctaaccatca gaaccggggg aaggaccatc
3241 accacgggtg ccccatcac gtactccacc tatggcaagt ttcttgccga cgggtggttg
3301 tctgggggcg cctatgacat cataatatgt gatgagtgcc actcaactga ctcgaccact
3361 atctctggca ctggcacagt cctggaccaa gcggagacgg ctggagcgcg actcgtcgtg
3421 ctgcgccacc ctaccctcc gggtagcgtc accgtgccac atccaaacat cgaggagggtg
3481 gctctgtcca gcactggaga aatccccctt tatggcaaag ccatcccat cgagaccatc
3541 aaggggggga ggcacctcat ttcttgccat tccaagaaga aatgtgatga gctcgccgcg
3601 aagctgtccg gcctcggact caatgctgta gcatattacc ggggccttga tgtatccgtc
3661 ataccaacta gcggagacgt cattgtcgtg gcaacggacg ctctaatac gggctttacc
3721 ggcgatttcg actcagtgat cgactgcaat acatgtgtca ccagacagt cgacttcagc
3781 ctggaccgca ccttcaccat tgagacgacg accgtgccac aagacgcggg gtacgcgtcg
3841 cagcggcgag gcaggactgg taggggcagg atgggcattt acaggtttgt gactccagga
3901 gaacggccct cgggcattgt cgattcctcg gttctgtgag agtgctatga cgcgggctgt
3961 gcttggtacg agctcacgcc cgcgcgagac tcagttagggt tgcgggctta cctaaccaca
4021 ccagggttgc ccgtctgcca ggaccatctg gagttctggg agagcgtctt tacaggcctc
4081 acccacatag acgcccattt ctgtcccag actaagcagg caggagacaa ctccccctac
4141 ctggtagcat accaggctac ggtgtgcgcc agggctcagg ctccacctcc atcgtgggac
4201 caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gcccttctg
4261 tataggctgg gagccgttca aaacgaggtt actaccacac accccataac caaatacatc
4321 atggcatgca tgcggctga cctggaggtc gtcacgagca cctgggtgtg ggtaggcgga
4381 gtccatagcag ctctggccgc gtattgcctg acaacaggca gcgtggctat tgcgggcagg
4441 atcatcttgt ccggaaagcc ggccatcatt cccgacaggg aagtccttta ccgggagttc
4501 gatgagatgg aagagtgcgc ctcacacctc ccttacatcg aacagggaat gcagctcgcc
4561 gaacaattca aacagaaggc aatcgggttg ctgcaaacag ccaccaagca agcggaggct
4621 gctgctcccg tgggtggaat caagtggcgg accctcgaag ccttctgggc gaagcatatg
4681 tggaaatttca tcagcgggat acaatattta gcaggcttgt cactctgcc tggcaacccc
4741 gcgatagcat cactgatggc attcacagcc tctatcacca gcccgctcac caccacacat
4801 accctcctgt ttaacatcct ggggggatgg gtggccgccc aacttgctcc tccagcgct
4861 gctctgctt tcgtaggcgc cggcatcgct ggagcggctg tggcagcat aggcctggg
4921 aagggtcttg tggatatttt ggcaggttat ggagcagggg tggcagcgc gctcgtggc
4981 ttaagggtca tgagcggcga gatgccctcc accgaggacc tggttaacct actcctgct
5041 atcctctccc ctggcgcctt agtcgtcggg gtcgtgtgag cagcgatact gcgtcggcac
5101 gtgggcccag gggagggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgcgg
5161 ggtaaccacg tctcccccac gcactatgtg cctgagagcg acgctgcagc acgtgtcact
5221 cagatcctct ctagtcttac catcactcag ctgctgaaga ggcttcacca gtggatcaac
5281 gaggactgct ccacgccatg ctccggctcg tggctaagag atgtttggga ttggatatgc
5341 acgggttga ctgatttcaa gacctggctc cagtccaagc tctgcccgcg attgcccggg
5401 gtccccctt tctcatgtca acgtgggtac aaggaggtct ggcggggcga cggcatcatg
5461 caaaccacct gcccatgtgg agcacagatc accggacatg tgaaaaacgg ttccatgagg
5521 atcgtggggc ctaggacctg tagtaacacg tggcatggaa cattccccat taacgcgtac
5581 accacggggc cctgcacgcc ctccccggcg ccaaattatt ctaggcgct gtggcgggtg
5641 gctgctgagg agtacgtgga ggttacgcgg gtgggggatt tccactacgt gacgggcatg
5701 accactgaca acgtaaagtg cccgtgtcag gttccggccc ccgaattctt cacagaagtg
5761 gatgggggtg ggttgacag gtacgtcca gcgtgcaaac ccctcctacg ggaggaggtc
5821 acattcttgg tcgggtccta tcaatatcct gttgggtcac agtccccatg cgagcccgaa
5881 caggacgtag cagtgtcac ttccatgctc accgacctt cccacattac ggcggagacg
5941 gctaagcgta gggtggccag gggatctccc ccctccttgg ccagctcatc agctagccag
6001 ctgtctgcgc ctctcttgaa ggcaacatgc actaccgctc atgactcccc ggacgctgac
6061 ctcatcgagg ccaacctcct gtggcggcag gagatgggag ggaacatcac ccgctggag
```

6121 tcagaaaata aggtagtaat tttggactct ttcgagccgc tccaagcggga ggaggatgag
6181 aggggaagtat ccgttcgggc ggagatcctg cggagggtcca ggaaattccc tcgagcgatg
6241 cccatatggg cacgcccggg ttacaacctt ccactgttag agtcctggaa ggaccgggac
6301 tacgtccctc cagtgggtaca cgggtgtcca ttgccgcctg ccaaggcccc tccgatacca
6361 cctccacgga ggaagaggac ggttgtcctg tcagaatcta ccgtgtcttc tgccttggcg
6421 gagctcgcca caaagacctt cggcagctcc gaatcgctcg ccgtcgacag cggcacggca
6481 acggcctctc ctgaccagcc ctccgacgac ggcgacgcgg gatccgacgt tgagtctgac
6541 tcctccatgc ccccccttga gggggagccg ggggatcccg atctcagcga cgggtcttgg
6601 tctaccgtaa gcgaggaggc tagtgaggac gtcgtctgct gctcgatgtc ctacacatgg
6661 acaggcgccc tgatcacgcc atgcgctgcg gaggaaacca agctgccccat caatgcactg
6721 agcaactctt tgctccgtca ccacaacttg gtctatgcta caacatctcg cagcgcaagc
6781 ctgcggcaga agaagggtcac ctttgacaga ctgcagggtcc tggacgacca ctaccgggac
6841 gtgctcaagg agatgaaggc gaaggcgctc acagttaagg ctaaacttct atccgtggag
6901 gaagcctgta agctgacgcc ccacattctg gccagatcta aatttggcta tggggcaaag
6961 gacgtccgga acctatccag caaggccgtt aaccacatcc gctccgtgtg gaaggacttg
7021 ctggaagaca ctgagacacc aattgacacc accatcatgg caaaaaatga ggttttctgc
7081 gtccaaccag agaagggggg ccgcaagcca gctcgcctta tcgtattccc agatttgggg
7141 gttcgtgtgt gcgagaaaat ggccctttac gatgtggtct ccacctccc tcaggccgtg
7201 atgggctctt catacggatt ccaatactct cctggacagc gggtcgagtt cctgggtgaat
7261 gcctggaaag cgaagaaatg ccttatgggc ttcgcatatg acaccgcctg ttttgactca
7321 acggtcactg agaatgacat ccgtgttgag gagtcaatct accaatgttg tgacttggcc
7381 cccgaagcca gacaggccat aaggctgcctc acagagcggc tttacatcgg gggccccctg
7441 actaattcta aagggcagaa ctgcggtctat cgccggtgcc gcgcgagcgg tgtactgacg
7501 accagctgcg gtaataccct cacatgttac ttgaaggccg ctgcggcctg tcgagctgcg
7561 aagctccagg actgcacgat gctcgtatgc ggagacgacc ttgtcgttat ctgtgaaagc
7621 gcggggaccc aagaggacga ggcgagccta cgggccttca cggaggctat gactagatac
7681 tctgcccccc ctggggaccc gcccaaacca gaatacgact tggagttgat aacatcatgc
7741 tcctccaatg tgtcagtcgc gcacgatgca tctggcaaaa ggggtgtacta tctcaccgt
7801 gacccccacca ccccccttgc gcgggtgcg tgggagacag ctagacacac tccagtcaat
7861 tcctggctag gcaacatcat catgtatgcg ccacacctgt gggcaaggat gatcctgatg
7921 actcatttct tctccatcct tctagctcag gaacaacttg aaaaagccct agattgtcag
7981 atctacgggg cctgttactc cattgagcca cttgacctac ctgagatcat tcaacgactc
8041 catggcctta gcgcattttc actccatagt tactctccag gtgagatcaa taggggtggct
8101 tcatgcctca ggaaacttgg ggtaccgccc ttgcgagctc ggagacatcg ggccagaagt
8161 gtccgcgcta ggctactgtc ccaggggggg agggctgcc aattgtggca gtacctcttc
8221 aactgggcag taaggaccaa gctcaaactc actccaatcc cggctgcgtc ccagttggat
8281 ttatccagct ggttcgttgc tgggtacagc gggggagaca tatatcacag cctgtctcgt
8341 gcccgacccc gctggttcat gtggtgccta ctactacttt ctgtaggggt aggcatctat
8401 ctactcccca accgatgaac ggggagctaa acactccagg ccaataggcc atcctgtttt
8461 tttccctttt tttttttctt tttttttttt tttttttttt tttttttttt ttctcctttt
8521 tttttcctct ttttttctt ttctttctt ttgtggctcc atcttagccc tagtcacggc
8581 tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc
8641 agatcaagt

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Mar 15 2004 07:22:16



Nucleotide

Entrez

PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

B_i

Search for

Display default Show 20 Send to File Clipboard Details

1: AJ242651. Hepatitis C virus...[gi:5441831]

Links

```

LOCUS       SSE242651               8637 bp    RNA            linear      SYN 11-MAY-2000
DEFINITION  Hepatitis C virus replicon I377/NS2-3'UTR.
ACCESSION   AJ242651
VERSION     AJ242651.1   GI:5441831
KEYWORDS    core-neo gene; NS2 gene; NS2 proteinase; NS3 gene; NS3
            proteinase/helicase; NS3/4A proteinase cofactor; NS4A gene; NS4B
            gene; NS4b protein; NS5a gene; NS5A phosphoprotein; NS5b gene; NS5B
            RNA dependant RNA polymerase; polyprotein.
SOURCE      Hepatitis C virus replicon I377/NS2-3'UTR
  ORGANISM  Hepatitis C virus replicon I377/NS2-3'UTR
            artificial sequences; vectors.
REFERENCE   1
  AUTHORS   Lohmann,V., Korner,F., Koch,J., Herian,U., Theilmann,L. and
            Bartenschlager,R.
  TITLE     Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell
            line
  JOURNAL   Science 285 (5424), 110-113 (1999)
  MEDLINE   99322193
  PUBMED    10390360
REFERENCE   2 (bases 1 to 8637)
  AUTHORS   Bartenschlager,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-MAY-1999) Bartenschlager R., Institute for Virology,
            Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67,
            55131 Mainz, GERMANY
FEATURES             Location/Qualifiers
     source            1..8637
                     /organism="Hepatitis C virus replicon I377/NS2-3'UTR"
                     /mol_type="other RNA"
                     /db_xref="taxon:95362"
                     /focus
                     /note="Neomycin-selectable bicistronic subgenomic
             hepatitis C virus replicon"
     source            join(1..341,1801..8406)
                     /organism="Hepatitis C virus"
                     /mol_type="other RNA"
                     /strain="1b"
                     /isolate="Con1"
                     /db_xref="taxon:11103"
                     /tissue_type="liver"
                     /country="Germany"
     source            1190..1800
                     /organism="Encephalomyocarditis virus"
                     /mol_type="other RNA"
                     /db_xref="taxon:12104"
     5'UTR          1..341
                     /evidence=experimental
     gene            342..1181
                     /gene="core-neo"
     CDS            342..1181

```



```
/gene="core-neo"
/function="selectable marker"
/codon_start=1
/product="core-neo fusion protein"
/protein_id="CAB46910.1"
/db_xref="GI:5441832"
/translation="MSTNPKPQRKTKGRAMIEQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQGRPVLFVKTDLSGALNELQDEAARLSWLATTGVPCAAALVDVVTEAGRDWLLLGEPVPGQDLLSSHLAPAEKVSIMADAMRRLHTLDPATCFPDHQAKHRIERARTRMEAGLVDQDDLDDEEHQGLAPAEFLARLKARMPDGEDLVVTHGDACLPNIMVENGRFSGFIDCGRLGVADRYQDIALATRDIAEELGGEWADRFLVLYGIAAPDSQRIAFYRLLEDFF"
misc_feature 1190..1800
              /note="internal ribosome entry site (IRES)"
CDS           1801..8406
              /codon_start=1
              /product="non-structural polyprotein"
              /protein_id="CAB46911.1"
              /db_xref="GI:5441833"
              /translation="MDREMAASCGGAVFVGLILLTSLPHYKFLARLIWWLQYFITRAEHLQVWIPPLNVRGGRDAVILLTCAIHPELIFTITKILLAILGLMVLQAGITKVPYFVRAHGLIRACMLVRKVAGGHYVQMALMKLAALTGTYYVDHLTPLRDWAHAGLRDLAVAVEPVVFSDMETKVIITWGADTAACGDIILGLPVSARRGREIHLGPADSLEGQGWRLLA  
PITAYSQQTRGLLGCIITSLTGRDRNQVEGEVQVVSTATQSFLATCVNGVCWTVYHGA  
GSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYLVTRHADVIPVR  
RRGDSRGSLLSPRPVSYLKSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVESH  
ETTMRSVPVFTDNSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVA  
ATLGFGAYMSKAHGIDPNIRTGVRTITTTGAPITYSTYKFLADGGCSGGAYDIIICDE  
CHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSSTGEIPF  
YGKAIP IETIKGGRHLIFCHSKKKCDELA AKLSGLGLNAVAYYRGLDVSVIP TSGDVI  
VVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGRT  
GRGRMG IYRFVTPGERPSGMFDSVLCECYDAGCAWYELTPAETSVRRLRAYLNTPGLP  
VCQDHLEFWESVFTGLTHIDAHFLSQTQKAGDNFPYLVAYQATVCARAQAPPPSWDQM  
WKCLIRL KPTLHGPTPLLYRLGAVQNEVTTHPITKYIMACMSADLEVVTSTWVLVGG  
VLAALAA YCLTTGSVVIVGRIILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQ  
LAEQFKQKAIGLLQTATKQAEAAAPVVESKWR TLEAFWAKHMWNFISGIQYLAGLSTL  
PGNPAIASLMAFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASAFVGAGIAGAAV  
GSI GLGKVLVDILAGYGAGVAGALVAFKVMSEMPSTEDLVNLLPAILSPGALVVGVV  
CAAILRRHVGPGE GAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTQILSSLTITQ  
LLKRLHQWINE DCSTPCSGSWLRD VWDWICTVLTDFKTWLQSKLLPRLPGVPFFSCQR  
GYKGVWRGDGIMQTTCPCGAQITGHVKNGSMRIVGPRTCSNTWHGTFP INAYTTGPCT  
PSPAPNYSRALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPQCVPAP EFFEVEVDGVR  
LHRYAPACKPL LREEVTFVLVGLNQYL VGSQLPCEPEPDVAVLTSMLTDP SHITAETAK  
RRLARGSPPSLASSSASQLSAPSLKATCTTRHDSPADLIEANLLWRQEMGGNITRVE  
SENKVVILDSFEPLQAEEDEREVSVP AEILRRSRKFPRAMPIWARPDYNPPLLESWKD  
PDYVPPVHGCPLPPAKAPP IPPRRKRRTVVLSESTVSSALAE LATKTFGSSESAVD  
SGTATASPDQPSDDGDAGSDVESYSSMPLEGE PGDPDLSDGSWSTVSEEASEDVVCC  
SMSYTWTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASLRQKKVTFDRLQ  
VLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFY GAKDVRNLSSKAV  
NHIRSVWKD LLED TETPIDTTIMAKNEVF CVQPEKGGRKPARLIVFPDLGVRVCEKMA  
LYDVVSTLPQAVMGSSYGFQYSPGQ RVEFLVNAWKAKCPMGFAYDTRCFDSTVTEND  
IRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCG  
NTLTCYLKAAAACRAAKLQDCTMLVCGDDL VVICESAGTQEDEASLRAFTEAMTRYSA  
PPGDPPKPEYDLELITSCSSNVSVAH DASGKRYYLTRDPTT PLARA AWETARHTPVN  
SWLGNIIMYAPTLWARMILMTHFFSILL AQEQLEKALDCQIYGACYSIEPLDLPQIIQ  
RLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRATCG  
KYLFWNAVRTKLKLTPIPAASQLDLSSWFVAGYSGGDIYHSLSRARPRWFMWCLLLLS  
VGVGIIYLLPNR"
mat_peptide 1801..2451
              /gene="NS2"
              /product="NS2 proteinase"
mat_peptide 2452..4344
              /gene="NS3"
              /product="NS3 proteinase/helicase"
```

mat_peptide 4345..4506
 /gene="NS4A"
 /product="NS3/4A proteinase cofactor"
mat_peptide 4507..5289
 /gene="NS4B"
 /product="NS4B protein"
mat_peptide 5290..6630
 /gene="NS5A"
 /product="NS5A phosphoprotein"
mat_peptide 6631..8403
 /gene="NS5B"
 /product="NS5B RNA dependant RNA polymerase"
gene 1801..2451
 /gene="NS2"
gene 2452..4344
 /gene="NS3"
gene 4345..4506
 /gene="NS4A"
gene 4507..5289
 /gene="NS4B"
gene 5290..6630
 /gene="NS5A"
gene 6631..8403
 /gene="NS5B"
3'UTR 8407..8637

ORIGIN

```

1 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg
61 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac
121 cccccctccc gggagagcca tagtgggtctg cggaaccggg gagtacaccg gaattgccag
181 gacgaccggg tcctttcttg gatcaaccgc ctcaatgcct ggagatttgg gcgtgcccc
241 gcgagactgc tagccgagta gtgttggtgc gcgaaaggcc ttgtggtact gcctgatagg
301 gtgcttgcca gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac
361 ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcagggttctc
421 cgcccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct
481 ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg
541 actgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca
601 cgacgggcgt tccttgccga gctgtgctcg acgttgtcac tgaagcggga agggactggc
661 tgctattggg cgaagtgcgc gggcaggatc tcctgtcatc tcaccttgc cctgccgaga
721 aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc
781 cattcgacca ccaagcgaac catcgcacgc agcgagcacg tactcggatg gaagccggctc
841 ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg
901 ccaggtcaa ggcgcgcacg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct
961 gcttgccgaa tatcatggtg gaaaatggcc gcttttcttg attcatcgac tgtggccggc
1021 tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc
1081 ttggcggcga atgggctgac cgcttccctg tgctttacgg tatcgccgct cccgattcgc
1141 agcgcacgc cttctatcgc cttcttgacg agttcttctg agtttaacaa gaccacaacg
1201 gtttccctct agcgggatca attccgcccc tctccctccc cccccctaa cgttactggc
1261 cgaagccgct tgggaataagg ccggtgtgcg ttgtctata tgttattttc caccatattg
1321 ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tctcttgac gagcattcct
1381 aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca
1441 gttcctcttg aagcttcttg aagacaacaa acgtctgtag cgaccctttg caggcagcgg
1501 aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct
1561 gcaaaaggcg cacaacccca gtgccacggt gtgagttgga tagttgtgga aagagtcaaa
1621 tggctctcct caagcgtatt caacaagggg ctgaaggatg ccagaaaggc acccatttgt
1681 atgggatctg atctggggcc tcggtgcaca tgctttacat gtgtttagtc gaggttaaaa
1741 aacgtctagg cccccgaac cacggggacg tggttttcct ttgaaaaaca cgataatacc
1801 atggaccggg agatggcagc atcgtgcgga ggcgcggttt tcgtaggtct gatactcttg
1861 accttgtcac cgcactataa gctgttcttc gctaggctca tatggtgggtt acaatatttt
1921 atcaccaggg ccgaggcaca cttgcaagtg tggatcccc cctcaacgt tcgggggggc
1981 cgcgatgccg tcactctcct cacgtgcgcg atccaccag agctaactct taccatcacc
2041 aaaatcttgc tcgccatact cgttccactc atggtgctcc aggcgtggtat aaccaaagtg
2101 ccgtacttgc tgcgcgcaca cgggctcatt cgtgcatgca tgctgggtgc gaagggtgct
2161 gggggctcatt atgtccaaat ggctctcatg aagttggccg cactgacagg tacgtacgtt
2221 tatgaccatc tcaccccaact gcgggactgg gccacgcgg gcctacgaga ccttgcgggtg

```

```
2281 gcagttgagc ccgtcgtctt ctctgatatg gagaccaagg ttatcacctg gggggcagac
2341 accgcggcgt gtggggacat catcttgggc ctgcccgtct ccgcccgcag ggggaggag
2401 atacatctgg gaccggcaga cagccttgaa gggcaggggt ggcgactcct gcgcctatt
2461 acggcctact cccaacagac gcgaggccta cttgggtgca tcatcactag cctcacaggc
2521 cgggacagga accaggtcga gggggagggt caagtggctt ccaccgcaac acaatctttc
2581 ctggcgacct gcgtcaatgg cgtgtgttgg actgtctatc atggtgccgg ctcaaagacc
2641 cttgccggcc caaagggccc aatcacccaa atgtacacca atgtggacca ggacctcgtc
2701 ggctggcaag cgccccccgg ggcgcggttc ttgacaccat gcacctgcgg cagctcggac
2761 ctttacttgg tcacgaggca tgccgatgtc attccgggtg gccggcgggg cgacagcagg
2821 gggagcctac tctccccag gcccgctctc tacttgaagg gctcttcggg cgggtccactg
2881 ctctgccctt cggggcacgc tgtgggcata tttcgggctg ccgtgtgcac ccgaggggtt
2941 gcgaaggcgg tggactttgt acccgctcag tctatggaaa ccactatgcg gtccccgggtc
3001 ttcacggaca actcgtcccc tccggccgta ccgcagacat tccagggtgg ccatctacac
3061 gcccctactg gtagcggcaa gagcactaag gtgccggctg cgtatgcagc ccaagggtat
3121 aaggtgcttg tctgaaccc gtccgctcgc gccaccctag gtttcggggc gtatatgtct
3181 aaggcacatg gtatcgaccc taacatcaga accggggtaa ggaccatcac cacgggtgcc
3241 cccatcacgt actccacctt tggcaagttt cttgccgacg gtggttgctc tgggggcgcc
3301 tatgacatca taatatgtga tgagtgccac tcaactgact cgaccactat cctgggcac
3361 ggcacagtc tggaccaagc ggagcggct ggagcgcgac tcgtcgtgct cgccaccgct
3421 acgcctccgg gatcgggtcac cgtgccacat ccaaacatcg aggagggtgg tctgtccagc
3481 actggagaaa tcccctttta tggcaaagcc atccccatcg agaccatcaa gggggggagg
3541 cacctcattt tctgccattc caagaagaaa tgtgatgagc tcgccgcgaa gctgtccggc
3601 ctcgactca atgtcttagc atattaccgg ggccttgatg tatccgtcat accaactagc
3661 ggagacgtca ttgtcgtagc aacggacgct ctaatgacgg gctttaccgg cgatttcgac
3721 tcagtgatcg actgcaatac atgtgtcacc cagacagtcg acttcagcct ggacccgacc
3781 ttcaccattg agacgacgac cgtgccacaa gacgcgggtg cacgctcgca gcggcgaggc
3841 aggactggta ggggcaggat gggcatttac aggtttgtga ctccaggaga acggccctcg
3901 ggcattgtcg attcctcggg tctgtgcgag tctatgacg cgggctgtgc ttggtacgag
3961 ctacgcgccg ccgagacctc agttagggtt cgggcttacc taaacacacc aggggtgccc
4021 gtctgccagg accatctgga gttctgggag agcgtcttta caggcctcac ccacatagac
4081 gcccatttct tgtccagac taagcaggca ggagacaact tccccacct ggtagcatc
4141 caggctacgg tgtgcgccag ggctcaggct ccacctccat cgtgggacca aatgtggaag
4201 tgtctcatac ggctaaagcc tacgctgcac gggccaacgc ccctgctgta taggctggga
4261 gccgttcaaa acgaggttac taccacacac ccataacca aatacatcat ggcattgatg
4321 tcggctgacc tggaggtcgt cacgagcacc tgggtgctgg taggcggagt cctagcagct
4381 ctggccgctg attgcctgac aacaggcagc gtggtcattg tgggcaggat catctgttc
4441 ggaaagccgg ccatcattcc cgacagggaa gtcccttaac gggagtctga tgagattgaa
4501 gagtgcgctt cacacctccc ttacatcgaa cagggaatgc agctcgccga acaattcaaa
4561 cagaaggcaa tcgggttgct gcaaacagcc accaagcaag cggaggctgc tgctcccggtg
4621 gtggaatcca agtggcggac cctcgaagcc ttctgggcga agcatatgtg gaatttcac
4681 agcgggatac aatatttagc aggcttgctc actctgcctg gcaacccgcg gatagcatca
4741 ctgatggcat tcacagcctc tatcaccagc ccgctcacca cccaacatac cctcctgttt
4801 aacatcctgg ggggatgggt ggccgcccac cttgctcctc ccagcgtgc tctgtcttc
4861 gtaggcgccg gcatcgctgg agcggctgtt ggcagcatag gccttgggaa ggtgctgtg
4921 gatattttgg caggttatgg agcaggggtg gcaggcgcgc tcgtggcctt taaggctcatg
4981 agcggcgaga tgccctccac cgaggacctg gttaacctac tccctgctat cctctccct
5041 ggcgccctag tcgtcggggg cgtgtgcgca gcgatactgc gtcggcacgt gggccagggtg
5101 gagggggctg tgcagtggat gaaccggctg atagcgttcg cttcgcgggg taaccacgtc
5161 tccccacgc actatgtgcc tgagagcgac gctgcagcac gtgtcactca gatcctctct
5221 agtcttacca tcactcagct gctgaagagg cttcaccagt ggatcaacga ggactgctcc
5281 acgcatgct ccggctcgtg gctaagagat gtttgggatt ggatatgcac ggtgttgact
5341 gatttcaaga cctggctcca gtccaagctc ctgccgcgat tgccgggagt cccctctctc
5401 tcatgtcaac gtgggtacaa gggagcttgg cggggcgacg gcatcatgca aaccacctgc
5461 ccatgtggag cacagatcac cgacatgtg aaaaacgggt ccatgaggat cgtggggcct
5521 aggacctgta gtaacacgtg gcatggaaca ttccccatta acgcgtacac cacgggcccc
5581 tgcacgcctt ccccgcgccc aaattattct agggcgctgt ggcgggtggc tgctgaggag
5641 tacgtggagg ttacgcgggt gggggatttc cactacgtga cgggcatgac cactgacaac
5701 gtaaagtgcc cgtgtcagggt tccggccccc gaattcttca cagaagtgga tgggggtcgg
5761 ttgcacagggt acgctccagc gtgcaaaccc ctctacggg aggaggtcac attcctggtc
5821 gggctcaate aatacctggt tgggtcacag ctccatgcg agccgaacc ggacgtagca
5881 gtgctcactt ccatgctcac cgacccctcc cacattacgg cggagacggc taacgtagg
5941 ctggccagggt gatctcccc ctccttggcc agctcatcag ctagccagct gtctgcgcct
6001 tccttgaagg caacatgcac taccgctcat gactccccgg acgctgacct catcgaggcc
6061 aacctcctgt ggcggcagga gatgggcggg aacatcacc gcgtggagtc agaaaataag
```

6121 gtagtaattt tggactcttt cgagccgctc caagcggagg aggatgagag ggaagtatcc
6181 gttccggcgg agatcctgcg gaggtccagg aaattccctc gagcgatgcc catatgggca
6241 cgcccgggatt acaacctctc actgttagag tcttggaagg acccggaacta cgtccctcca
6301 gtggtacacg ggtgtccatt gccgcctgcc aaggccccctc cgataccacc tccacggagg
6361 aagaggacgg ttgtcctgtc agaatactacc gtgtcttctg ccttggcggga gctcgccaca
6421 aagaccttcg gcagctccga atcgctggcc gtcgacagcg gcacggcaac ggctctctct
6481 gaccagccct ccgacgacgg cgacgcggga tccgacgttg agtcgtactc ctccatgccc
6541 ccccttgagg gggagccggg ggatcccgat ctcagcgacg ggtcttggtc taccgtaagc
6601 gaggaggcta gtgaggacgt cgtctgctgc tcgatgtcct acacatggac aggcgccttg
6661 atcacgccat gcgctgcgga ggaaaccaag ctgcccatac atgcactgag caactctttg
6721 ctccgtcacc acaacttggc ctatgctaca acatctcgca gcgcaagcct gcgcagaaag
6781 aaggtcacct ttgacagact gcaggctctg gacgaccact accgggacgt gctcaaggag
6841 atgaaggcga aggcgtccac agttaaggct aaacttctat ccgtggagga agcctgtaag
6901 ctgacgcccc cacattcggc cagatctaaa tttggctatg gggcaaagga cgtccggaac
6961 ctatccagca aggcggttaa ccacatccgc tccgtgtgga aggacttgct ggaagacact
7021 gagacaccaa ttgacaccac catcatggca aaaaatgagg ttttctgctg ccaaccagag
7081 aagggggggc gcaagccagc tcgccttata gtattcccag atttgggggt tcgtgtgtgc
7141 gagaaaatgg ccctttacga tgtggtctcc accctccctc aggcctgatg gggctcttca
7201 tacggattcc aatactctcc tggacagcgg gtcgagttcc tgggtgaatgc ctggaaagcg
7261 aagaaatgcc ctatgggctt cgcataatgc acccgctgtt ttgactcaac ggtcactgag
7321 aatgacatcc gtgttgagga gtcaatctac caatgttgtg acttggcccc cgaagccaga
7381 caggccataa ggtcgctcac agagcggctt tacatcgggg gccccctgac taattctaaa
7441 gggcagaact gcggttatcg ccggtgccgc gcgagcgggt tactgacgac cagctgcggt
7501 aataccctca catgttactt gaaggccgct gcggcctgtc gagctgcgaa gctccaggac
7561 tgcacgatgc tcgtatgcgg agacgacct gtctgttatc gtgaaagcgc ggggacccaa
7621 gaggacgagg cgagcctacg ggcccttcac gaggctatga ctagatactc tgccccccct
7681 ggggacccgc ccaaaccaga atacgacttg gagttgataa catcatgctc ctccaatgtg
7741 tcagtcgcgc acgatgcac tggcaaaagg gtgtactatc tcaccctga cccaccacc
7801 ccccttgccg gggctgcgtg ggagacagct agacacactc cagtcaattc ctggctaggc
7861 aacatcatca tgtatgcgcc caccttgtgg gcaaggatga tcctgatgac tcatttcttc
7921 tccatccttc tagctcagga acaacttgaa aaagccctag attgtcagat ctacggggcc
7981 tgttactcca ttgagccact tgacctacct cagatcattc aacgactcca tggccttagc
8041 gcattttcac tccatagtta ctctccaggt gagatcaata ggggtggcttc atgcctcagg
8101 aaacttgggg taccgcccct gcgagtcctg agacatcggg ccagaagtgt ccgcgctagg
8161 ctactgtccc agggggggag ggttgccact tgtggcaagt acctcttcaa ctgggcagta
8221 aggaccaagc tcaaactcac tccaatcccg gctgcgtccc agttggattt atccagctgg
8281 ttctgttctg gttacagcgg gggagacata tatcacagcc tgtctcgtgc ccgacccgc
8341 tggttcatgt ggtgcctact cctactttct gtaggggtag gcactctatc actcccaac
8401 cgatgaacgg ggagctaaac actccaggcc aataggccat cctgtttttt tccctttttt
8461 tttttctttt tttttttttt tttttttttt tttttttttt ctctttttt tttctctttt
8521 ttttctttt ctttctttt gtggctccat cttagcccta gtcacggcta gctgtgaaag
8581 gtccgtgagc cgcttgactg cagagagtgc tgatactggc ctctctgacg atcaagt

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Mar 15 2004 07:22:16